

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 14:00:24 ; Search time 163 Seconds
(without alignments)
64.323 Million cell updates/sec

Title: US-09-675-650-4

Perfect score: 20

Sequence: 1 gaggtaggaagatagaacg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4172979 seqs, 262114271 residues

Total number of hits satisfying chosen parameters: 8345958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA New:*

- 1: /cgn2_6/ptodata/2/pubpna/US08 NEW PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06 NEW PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US07 NEW PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US09 NEW PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US10 NEW PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US11 NEW PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US12 NEW PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US13 NEW PUB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US14 NEW PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US15 NEW PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	20	100.0	506	US-10-773-440A-7
2	16.8	84.0	789	Sequence 7, Appli
3	16.8	84.0	3466	Sequence 4593, Ap
4	16.4	82.0	1165	Sequence 51092, A
5	16.4	82.0	1931	Sequence 64341, A
6	15.8	79.0	2376	Sequence 28759, A
7	15.8	79.0	27032	Sequence 42464, A
8	15.8	79.0	260209	Sequence 13468, A
9	15.8	79.0	611587	Sequence 23, Appl
10	15.4	77.0	1861	Sequence 209, App
11	15.4	77.0	3065	Sequence 33082, A
12	15.4	77.0	645179	Sequence 40303, A
13	15.2	76.0	201	Sequence 13293, A
14	15.2	76.0	687	Sequence 14464, A
15	15.2	76.0	1400	Sequence 48039, A
16	15.2	76.0	1904	Sequence 29945, A
17	15.2	76.0	2465	Sequence 26121, A
18	15.2	76.0	40987	Sequence 1049, Ap
19	15.2	76.0	67088	Sequence 13503, A
20	15.2	76.0	96128	Sequence 13365, A
21	15.2	76.0	120697	Sequence 13197, A
22	15	75.0	19	Sequence 859687, A
23	15	75.0	19	Sequence 859687, A

24	15	75.0	25	7	US-11-121-849-608462	Sequence 608462, A
25	15	75.0	11131	6	US-10-240-708-27	Sequence 27, Appl
26	14.8	74.0	19	8	US-11-101-244-53147	Sequence 53147, A
27	14.8	74.0	19	8	US-11-101-244-53227	Sequence 53227, A
28	14.8	74.0	19	8	US-11-101-244-53247	Sequence 53247, A
29	14.8	74.0	19	8	US-11-101-244-53324	Sequence 53324, A
30	14.8	74.0	19	8	US-11-101-244-164004	Sequence 164004, A
31	14.8	74.0	19	8	US-11-101-244-164092	Sequence 164092, A
32	14.8	74.0	19	8	US-11-101-244-734019	Sequence 734019, A
33	14.8	74.0	19	9	US-11-083-784-53147	Sequence 53147, A
34	14.8	74.0	19	9	US-11-083-784-53227	Sequence 53227, A
35	14.8	74.0	19	9	US-11-083-784-53247	Sequence 53247, A
36	14.8	74.0	19	9	US-11-083-784-53324	Sequence 53324, A
37	14.8	74.0	19	9	US-11-083-784-164004	Sequence 164004, A
38	14.8	74.0	19	9	US-11-083-784-164092	Sequence 164092, A
39	14.8	74.0	19	9	US-11-083-784-734019	Sequence 734019, A
c 40	14.8	74.0	201	6	US-10-995-561-659331	Sequence 659331, A
41	14.8	74.0	600	6	US-10-750-185-36560	Sequence 36560, A
42	14.8	74.0	661	6	US-10-750-185-823	Sequence 823, App
c 43	14.8	74.0	977	7	US-10-750-185-53139	Sequence 53139, A
44	14.8	74.0	1271	6	US-11-147-492-9	Sequence 9, Appli
45	14.8	74.0			US-10-750-185-29548	Sequence 29548, A

ALIGNMENTS

RESULT 1
US-10-773-440A-7/c
; Sequence 7, Application US/10773440A
; Publication No. US20050282170A1
; GENERAL INFORMATION:
; APPLICANT: Fradet, Yves
; APPLICANT: Chypre, Camille
; APPLICANT: Garon, Lyson
; TITLE OF INVENTION: Method to Detect Prostate Cancer in a Sample
; FILE REFERENCE: 1619.0180001
; CURRENT APPLICATION NUMBER: US/10/773.440A
; CURRENT FILING DATE: 2004-02-09
; PRIOR APPLICATION NUMBER: 60/445,436
; PRIOR FILING DATE: 2003-02-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-773-440A-7

Query Match 100.0%; Score 20; DB 6; Length 506;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GAGTAGGAAGGATAGAAACG 20
Db 111 GAGTAGGAAGGATAGAAACG 92

RESULT 2
US-10-467-657-4593/c
; Sequence 4593, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467.657
; CURRENT FILING DATE: 2003-08-11

```
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4593
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4593

Query Match      84.0%; Score 16.8; DB 6; Length 789;
Best Local Similarity 90.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAACG 20
Db 148 GAGTAGGAATCAAGAAACG 129

RESULT 3
US-10-750-185-51092
; Sequence 51092, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51092
; LENGTH: 3466
; TYPE: DNA
; ORGANISM: Bovine 19866880511595
US-10-750-185-51092

Query Match      84.0%; Score 16.8; DB 6; Length 3466;
Best Local Similarity 90.0%; Pred. No. 50;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAACG 20
Db 2449 GAGTAAGAGGATACAAACG 2468

RESULT 4
US-10-750-185-64341
; Sequence 64341, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922

; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4593
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4593

Query Match      84.0%; Score 16.8; DB 6; Length 789;
Best Local Similarity 90.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAACG 20
Db 148 GAGTAGGAATCAAGAAACG 129

RESULT 3
US-10-750-185-51092
; Sequence 51092, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51092
; LENGTH: 3466
; TYPE: DNA
; ORGANISM: Bovine 19866880511595
US-10-750-185-51092

Query Match      84.0%; Score 16.8; DB 6; Length 3466;
Best Local Similarity 90.0%; Pred. No. 50;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAACG 20
Db 2449 GAGTAAGAGGATACAAACG 2468

RESULT 4
US-10-750-185-64341
; Sequence 64341, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4593
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Bovine 19866881067804
US-10-750-185-64341

Query Match      82.0%; Score 16.4; DB 6; Length 1165;
Best Local Similarity 94.4%; Pred. No. 71;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAACG 18
Db 689 GAGTTGGAAGGATAGAAACG 706

RESULT 5
US-10-750-185-28759/c
; Sequence 28759, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28759
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Bovine 19866881378846
US-10-750-185-28759

Query Match      82.0%; Score 16.4; DB 6; Length 1931;
Best Local Similarity 94.4%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAACG 18
Db 1314 GAGAAGGAAGGATAGAAACG 1297

RESULT 6
US-10-750-185-42464/c
; Sequence 42464, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42464
; LENGTH: 2376
```

```

; TYPE: DNA
; ORGANISM: Bovine 19866880867458
; US-10-750-185-42464

Query Match      79.0%; Score 15.8; DB 6; Length 2376;
Best Local Similarity 89.5%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAAC 19
   |||||
Db 1437 GAGTAGGAAGGATAGAAAC 1419

RESULT 7
US-10-995-561-13468
; Sequence 13468, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13468
; LENGTH: 27032
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-13468

Query Match      79.0%; Score 15.8; DB 6; Length 27032;
Best Local Similarity 89.5%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAAC 19
   |||||
Db 14284 GAGAAGGAGGATAGAAAC 14302

RESULT 8
US-10-933-025-23
; Sequence 23, Application US/10933025
; Publication No. US20050265987A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN, STEVEN
; APPLICANT: HEMMERICH, STEFAN
; APPLICANT: TOMITA, MEGUMI
; TITLE OF INVENTION: Sulfotransferases and methods of use
; TITLE OF INVENTION: thereof
; FILE REFERENCE: UCAL-230CON
; CURRENT APPLICATION NUMBER: US/10/933,025
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 260209
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(260209)
; OTHER INFORMATION: n = A,T,C or G
; US-10-933-025-23

Query Match      79.0%; Score 15.8; DB 6; Length 260209;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAAC 19
   |||||
Db 64389 GATTAGGAGGATAGAAAC 64407

RESULT 9
US-11-117-187-209/c
; Sequence 209, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 209
; LENGTH: 611587
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-11-117-187-209

Query Match      79.0%; Score 15.8; DB 7; Length 611587;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAAC 19
   |||||
Db 476559 GAGAAGGAGGATAGAAAC 476541

RESULT 10
US-10-750-185-33082/c
; Sequence 33082, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33082
; LENGTH: 1861
; TYPE: DNA
; ORGANISM: Bovine 19866881295080
; US-10-750-185-33082

Query Match      77.0%; Score 15.4; DB 6; Length 1861;
Best Local Similarity 94.1%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 TAGGAAGGATAGAAACG 20
   |||||

```

```
Db      1352 TGGGAAGGATAGAAACG 1336
;
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14464
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-14464

Query Match      76.0%; Score 15.2; DB 6; Length 201;
Best Local Similarity 85.0%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GAGTAGGAAGGATAGAAACG 20
Db      62 GAGTGGGAAGGATTGAGACG 43

RESULT 14
US-10-750-185-48039/c
; Sequence 48039, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 48039
; LENGTH: 687
; TYPE: DNA
; ORGANISM: Bovine 19866881734474
US-10-750-185-48039

Query Match      76.0%; Score 15.2; DB 6; Length 687;
Best Local Similarity 85.0%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GAGTAGGAAGGATAGAAACG 20
Db      74 GAGTGGGAAGGTTAGAGACG 55

RESULT 15
US-10-750-185-29945
; Sequence 29945, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
;

Db      547755 GAGAAGGAAGGATAGAA 547771

RESULT 13
US-10-995-561-14464/c
; Sequence 14464, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13293
; LENGTH: 645179
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13293

Query Match      77.0%; Score 15.4; DB 6; Length 645179;
Best Local Similarity 94.1%; Pred. No. 2.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GAGTAGGAAGGATAGAA 17
Db      547755 GAGAAGGAAGGATAGAA 547771

RESULT 11
US-10-750-185-40303/c
; Sequence 40303, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 40303
; LENGTH: 3065
; TYPE: DNA
; ORGANISM: Bovine 19866880368159
US-10-750-185-40303

Query Match      77.0%; Score 15.4; DB 6; Length 3065;
Best Local Similarity 94.1%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGTAGGAAGGATAGAAA 18
Db      455 AGTAGGAAGGAAGAAA 439

RESULT 12
US-10-995-561-13293
; Sequence 13293, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13293
; LENGTH: 645179
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13293

Query Match      77.0%; Score 15.4; DB 6; Length 645179;
Best Local Similarity 94.1%; Pred. No. 2.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GAGTAGGAAGGATAGAA 17
Db      547755 GAGAAGGAAGGATAGAA 547771

RESULT 13
US-10-995-561-14464/c
; Sequence 14464, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
```

```
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 29945
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Bovine 19866880773646
US-10-750-185-29945

Query Match      76.0%; Score 15.2; DB 6; Length 1400;
Best Local Similarity 85.0%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 GAGTAGGAAGGATAGAAACG 20
      ||| ||||| ||||| |||||
Db      980 GAGAAGGAAGGAAGAAAG 999
```

Search completed: January 2, 2006, 15:23:18
Job time : 165 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 13:57:00 ; Search time 2130 Seconds
(without alignments)
533.741 Million cell updates/sec

Title: US-09-675-650-4

Perfect score: 20

Sequence: 1 gaggtagaagatagaacg 20

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	6	AX107860 Sequence
2	20	100.0	506	6	CQ855951 Sequence
3	20	100.0	506	6	AX107857 Sequence
4	20	100.0	5435	8	AF103908 Homo sapi
5	20	100.0	128480	8	AL359314 Human DNA
6	19	95.0	237234	14	AC141860 Gallus ga
7	18	90.0	130365	14	AC150126 Gallus ga
8	18	90.0	200558	14	AC150067 Gallus ga
9	18	90.0	262093	14	AC150039 Gallus ga
10	17.4	87.0	58609	14	AC142494 Rattus no
11	17.4	87.0	130244	14	AC013518 Homo sapi
12	17.4	87.0	130336	8	HS17K7
13	17.4	87.0	153206	9	AC102304
14	17.4	87.0	153305	14	AC139606
15	17.4	87.0	162197	9	AC139333
16	17.4	87.0	163280	14	AC128280
17	17.4	87.0	166525	8	AC022413
18	17.4	87.0	170028	5	CR792439 Zebrafish

19	17.4	87.0	175512	9	AC091536
20	17.4	87.0	176287	9	AL513347
21	17.4	87.0	176708	14	AC027548
22	17.4	87.0	184877	8	AC087189
23	17.4	87.0	186983	14	AC137795
24	17.4	87.0	196840	14	AC023255
25	17.4	87.0	197659	8	AC009245
26	17.4	87.0	197882	14	AC117056
27	17.4	87.0	209731	8	AC145908
28	17.4	87.0	226910	14	AC141314
29	17.4	87.0	233262	14	AC097288
30	17.4	87.0	235892	14	AC135536
31	17.4	87.0	237837	14	AC097987
32	17.4	87.0	240068	14	AC121019
33	17.4	87.0	244595	14	AC095144
34	17.4	87.0	246869	14	AC128974
35	17.4	87.0	269708	14	AC097787
36	17	85.0	132830	14	AC148864
37	17	85.0	134184	14	AC150028
38	17	85.0	134826	8	AL158210
39	17	85.0	217627	14	AC163546
40	17	85.0	304018	14	AC161133
41	16.8	84.0	792	6	A96152
42	16.8	84.0	4456	9	BC053089
43	16.8	84.0	56327	14	AC166940
44	16.8	84.0	65158	14	AC100429
45	16.8	84.0	90907	8	AL513543

ALIGNMENTS

RESULT 1
LOCUS AX107860 20 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 4 from Patent WO0123550.
ACCESSION AX107860
VERSION AX107860.1 GI:13923251
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Busse, U., Chypre, C. and Fradet, Y.
TITLE Pca3 messenger rna species in benign and malignant prostate tissues
JOURNAL Patent: WO 0123550-A 4 05-APR-2001;
Diagnocure Inc. (CA)

FEATURES
source
1..20
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="oligonucleotide"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATAGAAACG 20
|||||
Db 1 GAGTAGGAGGATAGAAACG 20

RESULT 2
LOCUS CQ855951/c 506 bp DNA linear PAT 31-AUG-2004
DEFINITION Sequence 7 from Patent WO2004070056.
ACCESSION CQ855951
VERSION CQ855951.1 GI:51850785
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Fradet, Y., Chypre, C., Piche, L. and Garon, G.

TITLE Method to detect prostate cancer in a sample

JOURNAL Patent: WO 2004/0056-A 7 19-AUG-2004; Diagnocure Inc. (CA)

FEATURES Location/Qualifiers

source 1..506

 /organism="Homo sapiens"

 /mol_type="unassigned DNA"

 /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 506;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20

Db 111 GAGTAGGAAGGATAGAAACG 92

RESULT 3

AX107857/c

LOCUS AX107857 506 bp DNA linear PAT 30-APR-2001

DEFINITION Sequence 1 from Patent WO0123550.

ACCESSION AX107857

VERSION AX107857.1 GI:13923249

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Busse, U., Chypre, C. and Fradet, Y.

TITLE Pca3 messenger rna species in benign and malignant prostate tissues

JOURNAL Patent: WO 0123550-A 1 05-APR-2001; Diagnocure Inc. (CA)

FEATURES Location/Qualifiers

source 1..506

 /organism="Homo sapiens"

 /mol_type="unassigned DNA"

 /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 506;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20

Db 111 GAGTAGGAAGGATAGAAACG 92

RESULT 4

AF103908/c

LOCUS AF103908 5435 bp DNA linear PRI 14-AUG-2000

DEFINITION Homo sapiens non-coding RNA DD3 gene, exons 2, 3, and 4.

ACCESSION AF103908

VERSION AF103908.1 GI:6165974

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 5435)

AUTHORS Bussemakers, M.J., van Bokhoven, A., Verhaegh, G.W., Smit, F.P., Karthaus, H.F., Schalken, J.A., Debruyne, F.M., Ru, N. and Isaacs, W.B.

TITLE DD3: a new prostate-specific gene, highly overexpressed in prostate

Cancer Res. 59 (23), 5975-5979 (1999)

10606244

2 (bases 1 to 5435)

Bussemakers, M.J.G., Van Bokhoven, A., Verhaegh, G.W., Smit, F.P., Karthaus, H.F.M., Schalken, J.A., Debruyne, F.M.J., Ru, N. and Isaacs, W.B.

Direct Submission

Submitted (28-OCT-1998) Urology Research Laboratory, University Hospital Nijmegen, P.O. Box 9101, Nijmegen 6500 HB, Netherlands

Location/Qualifiers

1..5435

 /organism="Homo sapiens"

 /mol_type="genomic DNA"

 /db_xref="taxon:9606"

 /chromosome="9"

 /map="9q21-q22"

 1..78

 /rpt_family="AluY"

 /rpt_type="dispersed"

533..697

 /feature="alternative exon present in 5% of cDNA clones"

 /number=2

 1035..1294

 /rpt_family="Alu"

 /rpt_type="dispersed"

 join(1571..1753,1981..5435)

 /product="non-coding RNA DD3"

 /note="transcript III"

 join(1571..1753,1981..3579)

 /product="non-coding RNA DD3"

 /note="transcript (major) II"

 join(1571..1753,1981..2517)

 /product="non-coding RNA DD3"

 /note="transcript I"

 1571..1753

 /number=3

 1981..5435

 /number=4

 2495..2499

 /note="transcript I"

 2517

 /note="transcript I"

 3553..3558

 /note="transcript (major) II"

 3579

 /note="transcript (major) II"

 5268..5423

 /note="LINE"

 /rpt_family="L1"

 /rpt_type="dispersed"

 5435

 /note="transcript III"

ORIGIN

Query Match 100.0%; Score 20; DB 8; Length 5435;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20

Db 1838 GAGTAGGAAGGATAGAAACG 1819

RESULT 5

AL359314/c

LOCUS AL359314 128480 bp DNA linear PRI 18-MAY-2005

DEFINITION Human DNA sequence from clone RP11-108L4 on chromosome 9 Contains the 5' end of the KIAA0367 gene and the 5' end of the PCA3 gene for prostate cancer antigen 3 (DD3), complete sequence.

ACCESSION AL359314

VERSION AL359314.14 GI:14132938

KEYWORDS HTG; DD3; KIAA0367; PCA3; prostate cancer.


```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE   1 (bases 1 to 128480)
AUTHORS    Leongamornlert,D.
TITLE      Direct Submission
JOURNAL    Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
COMMENT     On May 16, 2001 this sequence version replaced gi:13897374.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep/
            This sequence was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-108L4 is from the library RPC1-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBAC3.6
            ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vegas@sanger.ac.uk
            -----
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.

FEATURES             source
misc_feature         1..128480
                    /location:Qualifiers
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="9"
                    /clone="RP11-108L4"
                    /clone_lib="RPC1-11.1"
                    100
                    /note="Clone right end: RP11-214N16"
gene                 join(complement(35038..35164),
                    complement(AL161626.20:182145),
                    complement(AL161626.20:179205),
                    complement(AL161626.20:171461),
                    complement(AL161626.20:164908),
                    complement(AL161626.20:164254),
                    complement(AL161626.20:138098),
                    /gene="KIAA0367"
                    /locus_tag="RP11-214N16.3-004"
                    join(complement(35038..35164),
                    complement(AL161626.20:182145),
                    complement(AL161626.20:179205),
                    complement(AL161626.20:171461),
                    complement(AL161626.20:164908),
                    complement(AL161626.20:164254),
                    complement(AL161626.20:138098),
                    /gene="KIAA0367"
                    /locus_tag="RP11-214N16.3-004"
                    /product="KIAA0367"
                    /note="match: ESTs: BM834998.1 BU784204.1
                    match: CDNA: BC022571.1"
                    join(complement(35038..35161),
                    complement(AL161626.20:182145),
                    complement(AL161626.20:179205),
                    complement(AL161626.20:164908),
                    complement(AL161626.20:164254),
                    complement(AL161626.20:138098),
                    /gene="KIAA0367"
                    /locus_tag="RP11-214N16.3-004"
                    /product="KIAA0367"
                    /note="match: ESTs: B1552977.1 D61567.1"
                    join(complement(35038..35145),
                    complement(AL161626.20:182145),
                    complement(AL161626.20:179205),
                    complement(AL161626.20:171461),
                    complement(AL161626.20:164908),
                    complement(AL161626.20:164271),
                    complement(AL161626.20:163193),
                    complement(AL161626.20:155976),
                    /gene="KIAA0367"
                    /locus_tag="RP11-214N16.3-007"
                    /product="KIAA0367"
                    /note="match: ESTs: AL539646 D61567.1"
                    join(complement(35038..35131),
                    complement(AL161626.20:182145),
                    complement(AL161626.20:179205),
                    complement(AL161626.20:171219),
                    /gene="KIAA0367"
                    /locus_tag="RP11-214N16.3-001"
                    join(complement(35038..35131),
                    complement(AL161626.20:182145),
                    complement(AL161626.20:179205),
                    complement(AL161626.20:171219),
                    /gene="KIAA0367"
                    /locus_tag="RP11-214N16.3-001"
                    /product="KIAA0367"
                    /note="match: ESTs: BM310471.1 BM310778.1"
                    join(complement(35038..35105),
                    complement(AL161626.20:182145),
                    complement(AL161626.20:179205),
                    complement(AL161626.20:171461),
                    complement(AL161626.20:164908),
                    complement(AL161626.20:164271),
                    complement(AL161626.20:163193),
                    complement(AL161626.20:155976),
                    /gene="KIAA0367"
                    /locus_tag="RP11-214N16.3-001"
                    /standard_name="OTTHOMP000000021505"
                    /codon_start=1
                    /product="KIAA0367"
                    /protein_id="CAI10912.1"
                    /db_xref="GI:55859641"
                    /db_xref="InterPro:IPR001251"
                    /db_xref="UniProt/TREMBL:QST475"
                    /translation="MLKSCRSAPSPVRKPLILRLLSSEVGMNDIPFEGVLSPSA
                    ADMPEPNSLDNDTPRIKLTAPNLINSLDQSEGLSDDDNLDSPEDIDINVDL
                    DTPDADSFEYTGHDPTAKDQSESIPEYTABEERDNLRTVTVIGQEQRIDM
                    KVIEPVRVISHGGYGDGLNAITVFAACFLPDSSRADYHYVMENILFLYVISTLELMV
                    complement(AL161626.20:171461),
                    complement(AL161626.20:164908),
                    complement(AL161626.20:164254),
                    complement(AL161626.20:163193),
                    complement(AL161626.20:164271),
                    complement(AL161626.20:164908),
                    complement(AL161626.20:171461),
                    complement(AL161626.20:179205),
                    complement(AL161626.20:182145),
                    complement(AL161626.20:182145),
                    complement(AL161626.20:179205),
                    complement(AL161626.20:171461),
                    complement(AL161626.20:164908),
                    complement(AL161626.20:164271),
                    complement(AL161626.20:163193),
                    complement(AL161626.20:155976),
                    /gene="KIAA0367"
                    /locus_tag="RP11-214N16.3-001"
                    /standard_name="OTTHOMP000000021505"
                    /codon_start=1
                    /product="KIAA0367"
                    /protein_id="CAI10912.1"
                    /db_xref="GI:55859641"
                    /db_xref="InterPro:IPR001251"
                    /db_xref="UniProt/TREMBL:QST475"
                    /translation="MLKSCRSAPSPVRKPLILRLLSSEVGMNDIPFEGVLSPSA
                    ADMPEPNSLDNDTPRIKLTAPNLINSLDQSEGLSDDDNLDSPEDIDINVDL
                    DTPDADSFEYTGHDPTAKDQSESIPEYTABEERDNLRTVTVIGQEQRIDM
                    KVIEPVRVISHGGYGDGLNAITVFAACFLPDSSRADYHYVMENILFLYVISTLELMV

```

```

AEDYMYVLNGATPRRMPGLGWMKKCYQWIDRLRKLKLSFIIVHPSWFIITLAVT
RPFISKFSKIKYVNS"
  join(complement(35038..35105), .182224),
  complement(AL161626..20:182145..182224),
  complement(AL161626..20:179205..179405),
  complement(AL161626..20:171461..171634),
  complement(AL161626..20:168694..168702),
  complement(AL161626..20:164908..165010),
  complement(AL161626..20:164200..164271))
/genes="KIAA0367"
/locus tag="RP11-214N16.3-001"
/standard_name="OTTHUMP0000021506"
/codon_start=1
/product="KIAA0367"
/protein_id="CAI10913.1"
/db xref="GI:55859642"
/db xref="UniProt:IPRO01251"
/db xref="InterPro:IPRO01251"
/translations="MLKSCSRASFPSPVRKPLILRLRLSEDVGM DIPPEEGVLSPSA
ADMRPPPSLDLNDTHPRKIKLTAPNINLSLDQSGSLSDNLDSPDEIDINVDL
DTDFEADSFYTGHEPTANKSQQSESIPEYTAEEEREDNLRWTVVIGEQQRID
MKVIEPVRRVISHGGSGGYGGDLNLIIVFAACFLPDSSRADYHYVWENLFLVISTL
ELMVADYMYVLYLNGATP"
  join(complement(47725..53531), complement(46301..47063),
  complement(AL161626..20:182145..182224),
  complement(AL161626..20:179205..179405),
  complement(AL161626..20:171461..171631),
  complement(AL161626..20:168694..168702),
  complement(AL161626..20:164908..165010),
  complement(AL161626..20:164140..164271),
  complement(AL161626..20:163193..163279),
  complement(AL161626..20:155914..156012),
  complement(AL161626..20:141011..141322))
/genes="KIAA0367"
/locus tag="RP11-214N16.3-006"
  join(complement(47725..53531), complement(46301..47063),
  complement(AL161626..20:182145..182224),
  complement(AL161626..20:179205..179405),
  complement(AL161626..20:171461..171631),
  complement(AL161626..20:168694..168702),
  complement(AL161626..20:164908..165010),
  complement(AL161626..20:164140..164271),
  complement(AL161626..20:163193..163279),
  complement(AL161626..20:155914..156012),
  complement(AL161626..20:141011..141322))
/genes="KIAA0367"
/locus tag="RP11-214N16.3-006"

Query Match      100.0%; Score 20; DB 8; Length 128480;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATGAGAACG 20
|||||
Db 126936 GAGTAGGAGGATGAGAACG 126917

RESULT 6
AC141860
LOCUS
DEFINITION Gallus gallus clone TAM31-57D16, WORKING DRAFT SEQUENCE, 10 ordered
pieces.
ACCESSION AC141860
VERSION AC141860.2 GI:29501863
KEYWORDS HTG; HTGS_PHASE2; HTGS DRAFT.
SOURCE Gallus gallus
ORGANISM Gallus
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
  Phasianinae; Gallus
  1 (bases 1 to 237234)
REFERENCE
  Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
  Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: dft
Center clone name: 057D16

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8x average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 23553 bases at least Q40
Consensus quality: 235850 bases at least Q30
Insert size: 180000; agarose-fp
Insert size: 236334; sum-of-contigs
Quality coverage: 11.59x in Q20 bases; agarose-fp
Quality coverage: 8.83x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 12817: contig of 12817 bp in length
* 12918: gap of unknown length
* 85073: contig of 72156 bp in length
* 85173: gap of unknown length
* 88174: contig of 3541 bp in length
* 88715: gap of unknown length
* 113928: contig of 25114 bp in length
* 113929: gap of unknown length
* 114029: contig of 77639 bp in length
* 191667: gap of unknown length
* 191668: gap of unknown length

```

```

Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J. R., Karlins, E.,
Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q. L., Maduro, V. B.,
Margulies, E. H., Masello, C., Maskeri, B., McDowell, J.,
Paquirigan, C., Pearson, R., Portnoy, M. E., Prasad, A.,
Pardix-Dugue, N., Schandler, K., Schueler, M. G., Sison, C.,
Stantripop, S., Thomas, J. W., Thomas, P. J., Touchman, J. W., Vogt, J. L.,
Wetherby, K. D., Wiggins, L., Young, A. and Green, E. D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 237234)
Green, E. D.
Direct Submission
Submitted (19-MAR-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 237234)
Green, E. D.
Direct Submission
Submitted (03-APR-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Apr 3, 2003 this sequence version replaced gi:29124110.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: dft
Center clone name: 057D16

```

* 191768 200289: contig of 8522 bp in length
 * 200290 200389: gap of unknown length
 * 200390 211976: contig of 11587 bp in length
 * 211977 212076: gap of unknown length
 * 212077 214145: contig of 2069 bp in length
 * 214146 214245: gap of unknown length
 * 214246 224194: contig of 9949 bp in length
 * 224195 224294: gap of unknown length
 * 224295 237234: contig of 12940 bp in length.

FEATURES

source

1..237234
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9031"
 /clone="TAM31-57D16"
 /clone_lib="TAM31"

misc_feature

1..12817
 /note="assembly_fragment
 clone_end:T7
 vector_side:left"

gap

12818..12917

misc_feature

12918..85073
 /note="assembly_fragment"

gap

85074..85173

misc_feature

85174..88714
 /note="assembly_fragment"

gap

88715..88814

misc_feature

88815..113928
 /note="assembly_fragment"

gap

113929..114028

misc_feature

114029..131667
 /note="assembly_fragment"

gap

191668..191767

misc_feature

191768..200289
 /note="assembly_fragment"

gap

200290..200389

misc_feature

200390..211976
 /note="assembly_fragment"

gap

211977..212076

misc_feature

212077..214145
 /note="assembly_fragment"

gap

214146..214245

misc_feature

214246..224194
 /note="assembly_fragment"

gap

224195..224294

misc_feature

224295..237234
 /note="assembly_fragment
 clone_end:SP6
 vector_side:right"

ORIGIN

Query Match 95.0%; Score 19; DB 14; Length 237234;

Best Local Similarity 100.0%; Pred. No. 42;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTAGGAGGATAGAAAC 19

|||||

Db 227163 GAGTAGGAGGATAGAAAC 227181

RESULT 7

AC150126

LOCUS

DEFINITION Gallus gallus clone WAG-32P6, WORKING DRAFT SEQUENCE, 2 ordered

AC150126

LOCUS

DEFINITION Gallus gallus clone WAG-32P6, WORKING DRAFT SEQUENCE, 2 ordered

ACCESSION

AC150126

VERSION AC150126.1 GI:49533721

KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.

SOURCE

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 130365)

DOE Joint Genome Institute.

Unpublished

TITLE

2 (bases 1 to 130365)

DOE Joint Genome Institute.

REFERENCE

TITLE

JOURNAL

Direct Submission

Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA

94598-1698, USA

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: http://www.jgi.doe.gov

Project Information

Center Project Name: 2363925

Center clone name: Chk_32P6

Summary Statistics

Consensus quality: 129752 bases at least Q40

Consensus quality: 129883 bases at least Q30

Consensus quality: 130130 bases at least Q20

Estimated insert size: 138000; agarose-fp estimation

Estimated insert size: 130265; sum-of-contigs estimation

Quality coverage: 31.45 in Q20 bases; agarose-fp estimation

Quality coverage: 33.32 in Q20 bases; sum-of-contigs estimation.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 56305: contig of 56305 bp in length

* 56306 56405: gap of unknown length

* 56406 130365: contig of 73960 bp in length.

Location/Qualifiers

1..130365

/organism="Gallus gallus"

/mol_type="genomic DNA"

/db_xref="taxon:9031"

/clones="WAG-32P6"

/clone_lib="Texas A&M Wageningen Chicken BAC Library"

56306..56405

/estimated_length=unknown

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 1.15e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGTAGGAGGATAGAAAC 19

|||||

Db 27176 AGTAGGAGGATAGAAAC 27193

RESULT 8

AC150067

LOCUS

DEFINITION

AC150067

LOCUS

DEFINITION Gallus gallus clone CH261-73M16, WORKING DRAFT SEQUENCE, 3 ordered

200558 bp

DNA

linear

HTG 01-JUL-2004

```

AC150067
AC150067.1 GI:49533662
HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS
SOURCE
ORGANISM
Gallus gallus
Gallus gallus (chicken)

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
DOE Joint Genome Institute.
1 (bases 1 to 200558)
Unpublished
2 (bases 1 to 200558)
Unpublished
DOE Joint Genome Institute.
Direct Submission
Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
94598-1698, USA

COMMENT
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

-----
Project Information
Center Project Name: 2959111
Center clone name: JF2-73M16
-----
Summary Statistics
Consensus quality: 198774 bases at least Q40
Consensus quality: 199779 bases at least Q30
Consensus quality: 200130 bases at least Q20
Estimated insert size: 200000; agarose-fp estimation
Estimated insert size: 200358; sum-of-contigs estimation
Quality coverage: 10.32 in Q20 bases; sum-of-contigs estimation
Quality coverage: 10.3 in Q20 bases; agarose-fp estimation
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 13044: contig of 13044 bp in length
* 13045 13144: gap of unknown length
* 13145 117583: contig of 104439 bp in length
* 117584 117683: gap of unknown length
* 117684 200558: contig of 82875 bp in length.
FEATURES
Location/Qualifiers
source
1..200558
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/clone="CH261-73M16"
/clone_lib="CHORI-261 Chicken BAC Library"
13045..13144
gap
/estimated_length=unknown
117584..117683
gap
/estimated_length=unknown

ORIGIN
Query Match 90.0%; Score 18; DB 14; Length 200558;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAAC 19
|||||
Db 88846 AGTAGGAAGGATAGAAAC 88863

RESULT 9
AC150039/c

AC150039
AC150039.1 GI:49533634
HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS
SOURCE
ORGANISM
Gallus gallus
Gallus gallus (chicken)

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
DOE Joint Genome Institute.
1 (bases 1 to 262093)
Unpublished
2 (bases 1 to 262093)
Unpublished
DOE Joint Genome Institute.
Direct Submission
Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
94598-1698, USA

COMMENT
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

-----
Project Information
Center Project Name: 2975745
Center clone name: JF2-117B18
-----
Summary Statistics
Consensus quality: 261290 bases at least Q40
Consensus quality: 261579 bases at least Q30
Consensus quality: 261828 bases at least Q20
Estimated insert size: 235000; agarose-fp estimation
Estimated insert size: 261993; sum-of-contigs estimation
Quality coverage: 11.84 in Q20 bases; agarose-fp estimation
Quality coverage: 10.62 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 144289: contig of 144289 bp in length
* 144290 144389: gap of unknown length
* 144390 262093: contig of 117704 bp in length.
FEATURES
Location/Qualifiers
source
1..262093
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/clone="CH261-117B18"
/clone_lib="CHORI-261 Chicken BAC Library"
144290..144389
gap
/estimated_length=unknown

ORIGIN
Query Match 90.0%; Score 18; DB 14; Length 262093;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAAC 19
|||||
Db 173152 AGTAGGAAGGATAGAAAC 173135

RESULT 10
AC142494
LOCUS
AC142494 58609 bp DNA linear HTG 01-APR-2003

```

DEFINITION Rattus norvegicus clone CH230-164M14, *** SEQUENCING IN PROGRESS

*** 34 unordered pieces.

AC142494
VERSION AC142494.1 GI:29423811
KEYWORDS HTG; HTGS PHASE1.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciuromorphi; Muridae; Murinae; Rattus.

1 (bases 1 to 58609)

REFERENCE
AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Amin, A., Angiano, D., Allen, C., Allen, H., Albrooks, S., Ayodeji, M., Baca, E., Baden, H., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Barnes, M., Barnstead, M., Benham, F., Baldwin, D., Bandaru, K., Blair, J., Blankenship, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregregios, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, H., Louis, L., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Nuydasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwako, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindest, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Stetter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, J., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 58609)

JOURNAL
TITLE
REFERENCE
AUTHORS
JOURNAL

Submitted (01-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: KEPD
Center clone name: CH230-164M14
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 45901 bases at least Q40
Consensus quality: 50264 bases at least Q30
Consensus quality: 53610 bases at least Q20
Estimated insert size: 48084; sum-of-contigs estimation
Quality coverage: 1x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1284: contig of 1284 bp in length
* 1285: gap of unknown length
* 1384: contig of 1163 bp in length
* 1385: gap of unknown length
* 2547: contig of 1096 bp in length
* 2548: gap of unknown length
* 2647: contig of 1096 bp in length
* 2648: gap of unknown length
* 3743: contig of 1198 bp in length
* 3744: gap of unknown length
* 3844: contig of 1072 bp in length
* 5041: gap of unknown length
* 5042: contig of 1051 bp in length
* 5142: gap of unknown length
* 6214: contig of 1051 bp in length
* 6313: gap of unknown length
* 6314: contig of 1300 bp in length
* 7365: gap of unknown length
* 7465: gap of unknown length
* 8765: gap of unknown length
* 8864: contig of 1450 bp in length
* 10315: gap of unknown length
* 10316: contig of 1195 bp in length
* 10415: gap of unknown length
* 11609: contig of 1308 bp in length
* 11610: gap of unknown length
* 13018: contig of 1258 bp in length
* 13117: gap of unknown length
* 13175: contig of 1269 bp in length
* 14376: gap of unknown length
* 14476: contig of 1243 bp in length
* 15745: gap of unknown length
* 15844: contig of 1271 bp in length
* 17088: gap of unknown length
* 17188: contig of 1537 bp in length
* 18459: gap of unknown length
* 18559: contig of 1599 bp in length
* 20096: gap of unknown length
* 20196: contig of 1599 bp in length
* 21795: gap of unknown length
* 21895: contig of 1261 bp in length
* 23156: gap of unknown length
* 23256: contig of 1268 bp in length
* 24524: gap of unknown length
* 24624: contig of 1314 bp in length
* 25938: gap of unknown length
* 26038: contig of 1990 bp in length
* 28028: gap of unknown length
* 28128: contig of 1672 bp in length
* 29800: gap of unknown length
* 29900: contig of 1599 bp in length
* 31498: gap of unknown length
* 31599: contig of 2652 bp in length
* 34251: gap of unknown length
* 34351: contig of 1672 bp in length
* 36023: gap of unknown length
* 36123: contig of 1567 bp in length
* 37690: gap of unknown length
* 37790: contig of 2211 bp in length
* 40000: contig of 2211 bp in length

FEATURES		REFERENCE	
source		AUTHORS	
40001		TITLE	
40100: gap of unknown length		REFERENCE	
42123: contig of 2023 bp in length		AUTHORS	
42223: gap of unknown length		TITLE	
43470: contig of 1247 bp in length		JOURNAL	
43570: gap of unknown length		REFERENCE	
45004: contig of 1434 bp in length		AUTHORS	
45104: gap of unknown length		TITLE	
47787: contig of 2683 bp in length		REFERENCE	
47887: gap of unknown length		AUTHORS	
49173: contig of 1286 bp in length		TITLE	
49273: gap of unknown length		REFERENCE	
51640: contig of 2367 bp in length		AUTHORS	
51740: gap of unknown length		TITLE	
54702: contig of 2962 bp in length		REFERENCE	
54803: gap of unknown length		AUTHORS	
58603: contig of 3807 bp in length.		TITLE	
Location/Qualifiers		JOURNAL	
1. 58609		COMMENT	
/organism="Rattus norvegicus"		TITLE	
/mol_type="genomic DNA"		JOURNAL	
/db_xref="taxon:10116"		COMMENT	
/clones="CH230-164W14"		TITLE	
1285. .1384		JOURNAL	
/estimated_length=unknown		COMMENT	
2548. .2647		TITLE	
/estimated_length=unknown		JOURNAL	
3744. .3843		COMMENT	
/estimated_length=unknown		TITLE	
5042. .5141		JOURNAL	
/estimated_length=unknown		COMMENT	
6214. .6313		TITLE	
/estimated_length=unknown		JOURNAL	
7365. .7464		COMMENT	
/estimated_length=unknown		TITLE	
8765. .8864		JOURNAL	
/estimated_length=unknown		COMMENT	
10315. .10414		TITLE	
/estimated_length=unknown		JOURNAL	
11610. .11709		COMMENT	
/estimated_length=unknown		TITLE	
13018. .13117		JOURNAL	
/estimated_length=unknown		COMMENT	
14376. .14475		TITLE	
/estimated_length=unknown		JOURNAL	
15745. .15844		COMMENT	
/estimated_length=unknown		TITLE	
17088. .17187		JOURNAL	
/estimated_length=unknown		COMMENT	
18459. .18558		TITLE	
/estimated_length=unknown		JOURNAL	
20096. .20195		COMMENT	
/estimated_length=unknown		TITLE	
21795. .21894		JOURNAL	
Query Match		87.0%; Score 17.4; DB 14; Length 58609;	
Best Local Similarity		94.7%; Pred. No. 3.2e+02;	
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy 1 GAGTAGGAAGATAGAAAC 19			
Db 32519 GAGTAGGAAGGAAGAAAC 32537			
RESULT 11			
AC013518/c			
LOCUS			
DEFINITION			
AC013518			
ACCESSION			
AC013518.2			
VERSION			
AC013518.2			
KEYWORDS			
HTG; HTGS PHASE0.			
SOURCE			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
1 (bases 1 to 130244)			
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,			
Baldwin,J., Barna,N., Beckerly,K., Boguslavsky,L., Boukhgalter,B.,			
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,			
Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,			
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,			
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,			
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,			
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,			
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,			
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,			
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,			
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,			
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,			
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.			
Direct Submission			
Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome			
Research, 320 Charles Street, Cambridge, MA 02141, USA			
On Apr 11, 2000 this sequence version replaced gi:6403763.			
All repeats were identified using RepeatMasker:			
Smit, A.F.A. & Green, P. (1996-1997)			
http://ftp.genome.washington.edu/RW/RepeatMasker.html			
----- Genome Center			
Center: Whitehead Institute/ MIT Center for Genome Research			
Center code: WIBR			
Web site: http://www-seq.wi.mit.edu			
Contact: sequence_submissions@genome.wi.mit.edu			
----- Project Information			
Center project name: L1903			
Center clone name: 115_N_3			

* NOTE: This record contains 151 individual			
* sequencing reads that have not been assembled into			
* contigs. Runs of N are used to separate the reads			
* and the order in which they appear is completely			
* arbitrary. Low-pass sequence sampling is useful for			
* identifying clones that may be gene-rich and allows			
* overlap relationships among clones to be deduced			
* However, it should not be assumed that this clone			
* will be sequenced to completion. In the event that			
* the record is updated, the accession number will			
* be preserved.			
* 1			
* 728: contig of 728 bp in length			
* 729			
* 828: gap of 100 bp			
* 829			
* 1582: contig of 754 bp in length			
* 1583			
* 1682: gap of 100 bp			
* 1683			
* 2482: contig of 800 bp in length			
* 2483			
* 2582: gap of 100 bp			
* 2583			
* 3369: contig of 787 bp in length			
* 3370			
* 3469: gap of 100 bp			
* 3470			
* 4225: contig of 756 bp in length			
* 4226			
* 4325: gap of 100 bp			
* 4326			
* 5088: contig of 763 bp in length			
* 5089			
* 5189: contig of 762 bp in length			
* 5189			
* 6050: gap of 100 bp			
* 6051			
* 6818: contig of 768 bp in length			
* 6819			
* 6918: gap of 100 bp			
* 6919			
* 7686: contig of 768 bp in length			
* 7687			
* 8564: contig of 778 bp in length			
* 8565			
* 8664: gap of 100 bp			
* 8665			
* 9409: contig of 745 bp in length			
* 9410			
* 9509: gap of 100 bp			
* 9510			
* 10263: contig of 754 bp in length			
* 10264			
* 10363: gap of 100 bp			
* 10364			
* 11149: contig of 786 bp in length			

11150 11249: gap of 100 bp
11250 12019: contig of 770 bp in length
12019 12119: gap of 100 bp
12119 12860: contig of 741 bp in length
12860 12961: gap of 100 bp
12961 13712: contig of 752 bp in length
13712 13812: gap of 100 bp
13812 14574: contig of 762 bp in length
14574 14674: gap of 100 bp
14674 15443: contig of 769 bp in length
15443 15544: gap of 100 bp
15544 16289: contig of 746 bp in length
16289 16389: gap of 100 bp
16389 17170: contig of 781 bp in length
17170 17270: gap of 100 bp
17270 18032: contig of 762 bp in length
18032 18132: gap of 100 bp
18132 18903: contig of 771 bp in length
18903 19003: gap of 100 bp
19003 19795: contig of 792 bp in length
19795 19895: gap of 100 bp
19895 20654: contig of 759 bp in length
20654 20754: gap of 100 bp
20754 21498: contig of 744 bp in length
21498 21598: gap of 100 bp
21598 22366: contig of 768 bp in length
22366 22466: gap of 100 bp
22466 23232: contig of 786 bp in length
23232 23332: gap of 100 bp
23332 24101: contig of 769 bp in length
24101 24201: gap of 100 bp
24201 24912: contig of 710 bp in length
24912 25011: gap of 100 bp
25011 25793: contig of 782 bp in length
25793 25893: gap of 100 bp
25893 26674: contig of 781 bp in length
26674 26774: gap of 100 bp
26774 27524: contig of 750 bp in length
27524 27624: gap of 100 bp
27624 28401: contig of 777 bp in length
28401 28501: gap of 100 bp
28501 29257: contig of 756 bp in length
29257 29357: gap of 100 bp
29357 30136: contig of 779 bp in length
30136 30236: gap of 100 bp
30236 30998: contig of 762 bp in length
30998 31098: gap of 100 bp
31098 31858: contig of 760 bp in length
31858 32741: contig of 783 bp in length
32741 32841: gap of 100 bp
32841 33599: contig of 758 bp in length
33599 33699: gap of 100 bp
33699 34480: contig of 781 bp in length
34480 34580: gap of 100 bp
34580 35344: contig of 764 bp in length
35344 35444: gap of 100 bp
35444 36202: contig of 758 bp in length
36202 36302: gap of 100 bp
36302 37026: contig of 724 bp in length
37026 37126: gap of 100 bp
37126 37910: contig of 784 bp in length
37910 38010: gap of 100 bp
38010 38772: contig of 762 bp in length
38772 38872: gap of 100 bp
38872 39647: contig of 775 bp in length
39647 39747: gap of 100 bp
39747 40514: contig of 767 bp in length
40514 40614: gap of 100 bp
40614 41379: contig of 765 bp in length
41379 41479: gap of 100 bp
41479 42236: contig of 757 bp in length
42236 42336: gap of 100 bp
42336

42337 43080: contig of 744 bp in length
43080 43180: gap of 100 bp
43180 43962: contig of 782 bp in length
43962 44062: gap of 100 bp
44062 44837: contig of 775 bp in length
44837 44937: gap of 100 bp
44937 45723: contig of 786 bp in length
45723 45823: gap of 100 bp
45823 46607: contig of 784 bp in length
46607 46707: gap of 100 bp
46707 47493: contig of 786 bp in length
47493 47593: gap of 100 bp
47593 48327: contig of 734 bp in length
48327 48427: gap of 100 bp
48427 49171: contig of 744 bp in length
49171 49271: gap of 100 bp
49271 50018: contig of 747 bp in length
50018 50118: gap of 100 bp
50118 50886: contig of 768 bp in length
50886 50986: gap of 100 bp
50986 51755: contig of 769 bp in length
51755 51855: gap of 100 bp
51855 52633: contig of 778 bp in length
52633 52733: gap of 100 bp
52733 53508: contig of 775 bp in length
53508 53608: gap of 100 bp
53608 54371: contig of 763 bp in length
54371 54471: gap of 100 bp
54471 55232: contig of 761 bp in length
55232 55332: gap of 100 bp
55332 56121: contig of 789 bp in length
56121 56221: gap of 100 bp
56221 56972: contig of 751 bp in length
56972 57072: gap of 100 bp
57072 57848: contig of 776 bp in length
57848 57948: gap of 100 bp
57948 58698: contig of 750 bp in length
58698 58798: gap of 100 bp
58798 59568: contig of 770 bp in length
59568 60436: contig of 768 bp in length
60436 60536: gap of 100 bp
60536 61302: contig of 766 bp in length
61302 61402: gap of 100 bp
61402 62167: contig of 765 bp in length
62167 62267: gap of 100 bp
62267

Query Match 87.0%; Score 17.4; DB 14; Length 130244;
Best Local Similarity 94.7%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GAGTAGGAAGGATAGAAAC 19
Db 34364 GAGTAGGAAGGATAGAAAC 34346
|||||

RESULT 12
HS17K7/c

LOCUS
DEFINITION Human DNA sequence from clone RPI-17K7 on chromosome 11p13,
complete sequence.

ACCESSION AL035078.32 GI:5805137

VERSION HTG.

KEYWORDS Homo sapiens (human)

SOURCE Eukaryota; Metazoa;

ORGANISM Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 130336)

AUTHORS Martin, S.

TITLE Direct Submission

JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Aug 30, 1999 this sequence version replaced gi:5804969.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
RPI-17K7 is from the library RPCI-1 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES
 source
 1..130336
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="RZPD:RPCIF704K0717"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="p13"
 /clone="RPI-17K7"
 /clone_lib="RPCI-1"
 1
 /note="Clone_left_end: RPI-17K7"
 20999
 /note="Clone_right_end: RPI-65P5"
 84140..84150
 /note="1329 bases of Tn10 (J01829) removed here. This sequence represents the duplicated flanking sequence of the Tn10."
 130336
 /note="Clone_right_end: RPI-17K7"
ORIGIN
Query Match 87.0% Score 17.4; DB 8; Length 130336;
Best Local Similarity 94.7% Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATAGAAC 19
 |||||
Db 68214 GAGTTGGAAGGATAGAAC 68196

RESULT 13
AC102304
LOCUS
DEFINITION Mus musculus chromosome 7, clone RP24-403P14, complete sequence.
AC102304
ACCESSION
VERSION AC102304.6 GI:60678486
KEYWORDS HTG.
SOURCE
 Mus musculus (house mouse)
ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
 1 (bases 1 to 153206)
AUTHORS
 Birren,B., Nusbaum,C. and Lander,E.
TITLE
 Mus musculus chromosome 7, clone RP24-403P14
JOURNAL
 Unpublished
REFERENCE
 2 (bases 1 to 153206)
 Direct Submission

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Govette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,C., Murphy,T., Naylor,J., Nguyen,C., O'Connor,T., O'Donnell,P., O'Neill,D., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phukhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

3 (bases 1 to 153206)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Bouckgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phukhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

Submitted (08-FEB-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

4 (bases 1 to 153206)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Bouckgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phukhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falis, I., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Johnson, L., Howell, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Louissegh, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarunpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, N., Rose, R., Ruiz, S., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Sosa, J., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, B., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wiecezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 153305)
Worley, K.C.
Direct Submission
Submitted (07-FEB-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 153305)
Worley, K.C.
Direct Submission
Submitted (14-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Feb 8, 2003 this sequence version replaced gi:28269330.

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: KDQP
Center clone name: CH230-388D19

Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 140842 bases at least Q40
Consensus quality: 143414 bases at least Q30
Consensus quality: 145056 bases at least Q20
Estimated insert size: 141820; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

FEATURES
source
1. 153305
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-388D19"
1195..1294
/estimated_length=unknown
2359..2458
/estimated_length=unknown
gap
gap

* NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1194: contig of 1194 bp in length
* 1195: gap of unknown length
* 1295: contig of 1064 bp in length
* 2359: gap of unknown length
* 2459: contig of 1271 bp in length
* 3730: gap of unknown length
* 3829: contig of 1071 bp in length
* 4900: gap of unknown length
* 5001: contig of 1880 bp in length
* 6881: gap of unknown length
* 6981: contig of 1143 bp in length
* 8124: gap of unknown length
* 8223: contig of 1597 bp in length
* 8224: gap of unknown length
* 9821: contig of 1884 bp in length
* 9920: gap of unknown length
* 11805: contig of 1040 bp in length
* 12944: gap of unknown length
* 13044: gap of unknown length
* 13445: contig of 1353 bp in length
* 14398: gap of unknown length
* 14498: contig of 1846 bp in length
* 16344: gap of unknown length
* 16444: contig of 1414 bp in length
* 17858: gap of unknown length
* 17958: contig of 1708 bp in length
* 19666: gap of unknown length
* 19766: contig of 3657 bp in length
* 23423: gap of unknown length
* 23523: contig of 2761 bp in length
* 26284: gap of unknown length
* 30690: contig of 4307 bp in length
* 30691: gap of unknown length
* 30791: contig of 4935 bp in length
* 35726: gap of unknown length
* 35825: contig of 4955 bp in length
* 35826: gap of unknown length
* 40881: contig of 5699 bp in length
* 40881: gap of unknown length
* 46580: gap of unknown length
* 46680: contig of 7859 bp in length
* 54539: gap of unknown length
* 54639: contig of 11455 bp in length
* 66094: gap of unknown length
* 66193: contig of 10166 bp in length
* 76359: gap of unknown length
* 76459: contig of 9892 bp in length
* 86351: gap of unknown length
* 86352: contig of 10496 bp in length
* 86452: gap of unknown length
* 96947: contig of 10856 bp in length
* 97048: gap of unknown length
* 107904: contig of 8698 bp in length
* 108004: gap of unknown length
* 116702: contig of 15169 bp in length
* 116802: gap of unknown length
* 131971: contig of 21235 bp in length.
* 132071: Location/Qualifiers

REFERENCE	5 (bases 1 to 162197)
AUTHORS	Wilson,R.K.
TITLE	Direct Submission
JOURNAL	Submitted (28-JAN-2005) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT	On Sep 15, 2004 this sequence version replaced gi:51468472. ----- Genome Center Center: Washington University Genome Sequencing Center Center code: WUGSC Web site: http://genome.wustl.edu Contact: submissions@watson.wustl.edu ----- Summary Statistics Center project name: M_BB00083F12

gap	3730. .3829	/estimated_length=unknown
gap	4901. .5000	/estimated_length=unknown
gap	6881. .6980	/estimated_length=unknown
gap	8124. .8223	/estimated_length=unknown
gap	9821. .9920	/estimated_length=unknown
gap	11805. .11904	/estimated_length=unknown
gap	12945. .13044	/estimated_length=unknown
gap	14398. .14497	/estimated_length=unknown
gap	16344. .16443	/estimated_length=unknown
gap	17858. .17957	/estimated_length=unknown
gap	19666. .19765	/estimated_length=unknown
gap	23423. .23522	/estimated_length=unknown
gap	26284. .26383	/estimated_length=unknown
gap	30691. .30790	/estimated_length=unknown
gap	35726. .35825	/estimated_length=unknown
gap	40781. .40880	/estimated_length=unknown

Query Match 87.0%; Score 17.4; DB 14; Length 153305;
 Best Local Similarity 94.7%; Pred.No.3e+02; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	GAGTAGGAGGATAGAAAC	19
Db	24999	GAGCAGGAGGATAGAAAC	25017

RESULT 15	AC139333	Mus musculus BAC clone RP24-83F12 from 3, complete sequence.	linear	ROD 28-JAN-2005
LOCUS	AC139333	162197 bp	DNA	
DEFINITION	AC139333	Mus musculus BAC clone RP24-83F12 from 3, complete sequence.		
ACCESSION	AC139333	HTG.		
VERSION	AC139333.4	GI:52077879		
KEYWORDS				
SOURCE		Mus musculus (house mouse)		
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.		
REFERENCE		1 (bases 1 to 162197)		
AUTHORS		Kanchi,K., Haglund,K., Bielicki,L. and Meyer,R.		
TITLE		The sequence of Mus musculus BAC clone RP24-83F12		
JOURNAL		Unpublished (2001)		
REFERENCE		2 (bases 1 to 162197)		
AUTHORS		McPherson,J.D. and Waterston,R.H.		
TITLE		Direct Submission		
JOURNAL		Submitted (30-JAN-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE		3 (bases 1 to 162197)		
AUTHORS		Wilson,R.K.		
TITLE		Direct Submission		
JOURNAL		Submitted (20-AUG-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE		4 (bases 1 to 162197)		
AUTHORS		Wilson,R.K.		
TITLE		Direct Submission		
JOURNAL		Submitted (15-SEP-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		

Search completed: January 2, 2006, 14:41:11
Job time : 2135 secs

November 2005

Published Applications Nucleic Acid and Published Applications Amino Acid database searches now generate two sets of results each. The Published Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published Applications New databases, older published applications make up the Published Applications Main databases.

Searches run against Nucleic Acid Published Applications produce two sets of results, with the extensions **.nupbm** (Published Applications_NA_Main) and **.nupbn** (Published Applications_NA_New).
Searches run against Amino Acid Published Applications produce two sets of results, with the extensions **.rapbm** (Published Applications_AA_Main) and **.rapbn** (Published Applications_AA_New).

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 13:49:34 ; Search time 308 Seconds
(without alignments)
432.772 Million cell updates/sec

Title: US-09-675-650-4

Perfect score: 20

Sequence: 1 gactagaagatagaacg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq 21:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*
- 14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	4	Aaf30668
C 2	20	100.0	506	4	Aaf30666 Human dif
C 3	20	100.0	506	13	ADR27854
C 4	17.4	87.0	32323	11	ACN44098 Human pro
C 5	16.8	84.0	789	10	ABZ40002 N. gonorr
C 6	16.8	84.0	792	2	Aaz12053 Neisseria
C 7	16.8	84.0	792	14	ABE49028
C 8	16.8	84.0	2086	8	ACA41244
C 9	16.8	84.0	6147	3	Aaz53858
C 10	16.8	84.0	35871	4	AAK84974 Human imm
C 11	16.4	82.0	65	6	ABN55158 Mouse spl
C 12	16.4	82.0	102	12	ACH931983
C 13	16.4	82.0	316	6	ABN23941 Human ORF
C 14	16.4	82.0	347	3	ACH19786 Human sec
C 15	16.4	82.0	540	12	ACH80283 Human gen
C 16	15.8	79.0	214	12	ACH92461
C 17	15.8	79.0	300	2	Aaz13970 Human gen
C 18	15.8	79.0	304	5	ABV59234 Human pro
C 19	15.8	79.0	309	4	AAK57552 Human imm

C 20	15.8	79.0	372	14	AEB89287
C 21	15.8	79.0	506	12	ACH78427
C 22	15.8	79.0	548	12	ACH78761
C 23	15.8	79.0	579	13	ADx65454
C 24	15.8	79.0	598	4	AAK70588
C 25	15.8	79.0	598	4	AAK70587
C 26	15.8	79.0	1117	13	ADx27477
C 27	15.8	79.0	1367	13	ADx60978
C 28	15.8	79.0	1926	3	AAA08075
C 29	15.8	79.0	2690	10	ADG32755
C 30	15.8	79.0	2720	4	ADG32755
C 31	15.8	79.0	2844	3	AAC98219
C 32	15.8	79.0	3900	8	ABx63328
C 33	15.8	79.0	29559	14	AD213242
C 34	15.8	79.0	29560	13	ABD33580
C 35	15.8	79.0	38719	14	AEA61133
C 36	15.8	79.0	58723	11	ACN44178
C 37	15.8	79.0	86563	14	AEA61256
C 38	15.8	79.0	87563	14	AEA61200
C 39	15.8	79.0	110000	3	AAF22303_4
C 40	15.8	79.0	110000	13	ABD32627_0
C 41	15.8	79.0	171324	11	ACN43892
C 42	15.8	79.0	188888	6	ABQ75562
C 43	15.8	79.0	198161	6	ABK83564
C 44	15.8	79.0	198161	12	ADQ17348
C 45	15.8	79.0	198161	13	ADR52701

ALIGNMENTS

RESULT 1

AAF30668

ID AAF30668 standard; CDNA; 20 BP.

XX AC AAF30668;

DT 11-JUN-2001 (first entry)

DE Prostate cancer antigen 3 (PSA3) nucleic acid.

XX PCA3; prostate cancer; antigen; marker; differential expression;

KW diagnosis; therapy; human; ss.

XX Homo sapiens.

XX WO200123550-A2.

XX 05-APR-2001.

XX 29-SEP-2000; 2000WO-CA001154.

XX 29-SEP-1999; 99US-0156594P.

XX (DIAG-) DIAGNOCURE INC.

XX Busse U, Chypre C, Fradet Y;

XX WPI; 2001-258132/26.

XX Novel nucleic acid encoding differentially expressed prostate cancer antigen 3 mRNA containing additional sequence giving rise to long PCA3 mRNA, useful for diagnosis of mammal afflicted with prostate cancer.

XX Claim 6; Page 57; 60pp; English.

XX The present sequence comprises nucleotides from human prostate cancer antigen 3 (PCA3) long mRNA (see AAF30666). Claimed isolated nucleic acid molecules consist of 10-50 nucleotides which specifically hybridise to a differentially expressed long PCA3 mRNA, and are complementary to, or consist of, at least 10 consecutive nucleotides of the present sequence. Long PCA3 mRNA includes a 228 bp sequence, inserted between exons 3a and 4, which is absent in short PCA3 mRNA. Short PCA3 mRNA is associated with

CC prostate cancer. Long PCA3 RNA is associated with a non-malignant
CC prostatic state. Differential expression of these 2 PCA3 RNA species
CC provides protocols for the diagnosis of prostate disease, including a
CC method of diagnosing the presence or predisposition to develop prostate
CC cancer in a patient. Also provided are therapeutic methods that use a
CC nucleic acid encoding a differentially expressed PCA3 mRNA molecule, an
CC antisense sequence, a protein encoded by a differentially expressed PCA3
CC mRNA, or an antibody raised against such a protein

XX
SQ Sequence 20 BP; 9 A; 1 C; 8 G; 2 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATAGAAACG 20
|||||
Db 1 GAGTAGGAGGATAGAAACG 20
|||||

RESULT 2
AAF30666/c
ID AAF30666 standard; cDNA; 506 BP.
AC AAF30666;
XX
XX 11-JUN-2001 (first entry)
DT
XX Human differentially expressed PCA3 cDNA (long form).
DE
XX PCA3; prostate cancer; antigen; benign prostatic hyperplasia;
KW differential expression; diagnosis; gene therapy; chromosome 9; human;
KW ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH exon 1..26
FT /*tag= a
FT /number= 3
FT misc_RNA 27..254
FT /*tag= b
FT /note= "228 bp insertion"
FT exon 255..506
FT /*tag= c
FT /number= 4a
XX
XX WO200123550-A2.
XX
XX
XX 05-APR-2001.
XX
XX 29-SEP-2000; 2000WO-CA001154.
XX
XX 29-SEP-1999; 99US-0156594P.
XX
XX (DIAG-) DIAGNOCURE INC.
XX
XX Busse U, Chypre C, Pradet Y;
XX
XX WPI; 2001-258132/26.
XX
XX Novel nucleic acid encoding differentially expressed prostate cancer
XX antigen 3 mRNA containing additional sequence giving rise to long PCA3
XX mRNA, useful for diagnosis of mammal afflicted with prostate cancer.
XX
XX Claim 3(a); Fig 3; 60pp; English.
XX
XX The present sequence is that of an RT-PCR-amplified fragment of human
XX prostate cancer antigen 3 (PCA3) mRNA that includes an additional 228 bp
XX sequence, inserted between exons 3 and 4a. A 2nd amplified fragment (see
XX AAF30667) lacks this additional sequence. The additional sequence
XX interrupts the open reading frame of PCA3 protein, thereby yielding a
XX truncated PCA3 protein. The shorter form PCA3 RNA is associated with

CC prostate cancer whereas the longer form PCA3 RNA is associated with a non
CC -malignant prostatic state, such as benign prostatic hyperplasia. Based
CC on the differential expression of these 2 PCA3 RNA species, protocols for
CC the diagnosis of prostate disease are provided, including a method of
CC diagnosing the presence or predisposition to develop prostate cancer in a
CC patient. Also provided are therapeutic methods that use a nucleic acid
CC encoding a differentially expressed PCA3 mRNA molecule, an antisense
CC sequence, a protein encoded by a differentially expressed PCA3 mRNA, or
CC an antibody raised against such a protein

XX
SQ Sequence 506 BP; 132 A; 123 C; 108 G; 143 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 4; Length 506;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATAGAAACG 20
|||||
Db 111 GAGTAGGAGGATAGAAACG 92
|||||

RESULT 3
ADR27854/c
ID ADR27854 standard; DNA; 506 BP.
XX
XX ADR27854;
AC
XX
XX 04-NOV-2004 (first entry)
DT
XX Human prostate specific PCA3 sequence long fragment.
DE
XX Prostate cancer; PCA3; ds.
KW
XX
XX Homo sapiens.
XX
XX WO2004070056-A2.
XX
XX 19-AUG-2004.
XX
XX 09-FEB-2004; 2004WO-CA000170.
XX
XX 07-FEB-2003; 2003US-0445436P.
XX
XX (DIAG-) DIAGNOCURE INC.
XX
XX Pradet Y, Chypre C, Piche L, Garon G;
XX
XX WPI; 2004-615607/59.
XX
XX
XX Detecting prostate cancer comprises performing an in vitro nucleic acid
XX amplification assay on a sample using primers specific to PCA3 sequence
XX or prostate specific nucleic acid sequence.
XX
XX Disclosure; SEQ ID NO 7; 97pp; English.
XX
XX The invention relates to a novel method for detecting prostate cancer in
XX a human patient. The method comprises: performing an in vitro nucleic
XX acid amplification assay on a biological sample of the patient or its
XX extract, using a first primer pair, which is specific to a prostate
XX cancer specific PCA3 sequence and a second primer pair, which is specific
XX to a prostate specific nucleic acid sequence; and detecting the PCA3
XX sequence and the prostate specific nucleic acid sequence, where a
XX detection of the PCA3 nucleic acid sequence or its level correlates with
XX a risk of developing prostate cancer or to a presence of prostate cancer
XX in the patient, and where an absence of detection of the PCA3 nucleic
XX acid sequence or its lower level in the sample validates an absence of
XX prostate cancer or a lower risk of developing prostate cancer, when the
XX second prostate specific nucleic acid is detected. The invention further
XX comprises: a kit for assessing the presence of prostate cancer or the
XX risk of developing prostate cancer in a patient. The method and kits are
XX useful for detecting prostate cancer in a human patient and for
XX determining a predisposition, or presence of prostate cancer or
XX monitoring a progression of prostate cancer in a patient. This

CC polynucleotide sequence represents a fragment of the prostate specific
CC PCA3 gene sequence of the invention.
XX
SQ Sequence 506 BP; 132 A; 123 C; 108 G; 143 T; 0 U; 0 Other;
0; Gaps 0;

Query Match 100.0%; Score 20; DB 13; Length 506;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATAGAAACG 20
Db 111 GAGTAGGAGGATAGAAACG 92

RESULT 4
ACN44098/c
ID ACN44098 standard; DNA; 32323 BP.
XX
AC ACN44098;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG1641631.
XX
KW Cystostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
FN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
XX Claim 1; SEQ ID NO 376; Opp; English.

The present invention relates to novel DNA and protein sequences which
are associated with carcinomas. The sequences are useful for: (i) for
screening drug candidates; (ii) for screening of bioactive agent capable
of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
a bioactive agent capable of modulating the activity of CAP; (iv) for
evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
(x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
determining Carcinoma Associated (CA) gene copy number. In addition, the
CA genes are useful as DNA vaccines and the CAP are useful as markers of
carcinoma including lymphoma. The present sequence is one such CA coding
sequence. Note: This patent is an equivalent to basic patent
US2002182586A1, for which no sequence data was published

Sequence 32323 BP; 7372 A; 7775 C; 8641 G; 8364 T; 0 U; 171 Other;
0; Gaps 0;

Query Match 87.0%; Score 17.4; DB 11; Length 32323;
Best Local Similarity 94.7%; Pred. No. 3 8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATAGAAAC 19
Db 7698 GAGTAGGAGGATAGAAAC 7680

CC polynucleotide sequence represents a fragment of the prostate specific
CC PCA3 gene sequence of the invention.
XX
SQ Sequence 506 BP; 132 A; 123 C; 108 G; 143 T; 0 U; 0 Other;
0; Gaps 0;

Query Match 100.0%; Score 20; DB 13; Length 506;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATAGAAACG 20
Db 111 GAGTAGGAGGATAGAAACG 92

RESULT 5
ABZ40002/c
ID ABZ40002 standard; DNA; 789 BP.
XX
AC ABZ40002;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae nucleotide sequence SEQ ID 4593.
XX
KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX
OS Neisseria gonorrhoeae.
XX
FN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB002069.
XX
PR 12-FEB-2001; 2001GB-00003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Masighani V, Monaci E;
XX
DR WPI; 2003-058415/05.
XX
PT P-PSDB; ABP79032.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
medicament for treating or preventing N. gonorrhoeae infection.
XX
PS Disclosure; Page 525; 815pp; English.

The present invention relates to proteins from Neisseria gonorrhoeae.
Also disclosed are the nucleic acid molecules encoding the proteins and
antibodies that specifically bind to the proteins. The composition
comprising the protein, nucleic acid or antibody is useful for the
manufacture of a medicament for treating or preventing N. gonorrhoeae
infection, this may be in the form of a vaccine or gene therapy.
Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
molecules of the invention

Sequence 789 BP; 208 A; 184 C; 189 G; 208 T; 0 U; 0 Other;
0; Gaps 0;

Query Match 84.0%; Score 16.8; DB 10; Length 789;
Best Local Similarity 90.0%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATAGAAACG 20
Db 148 GAGTAGGAGGATAGAAACG 129

RESULT 6
AAZ12053/c
ID AAZ12053 standard; DNA; 792 BP.
XX
AC AAZ12053;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria gonorrhoeae complete ORF31 sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria gonorrhoeae.
XX
FN WO9924578-A2.
XX
PD 20-MAY-1999.
XX

```

PF 09-OCT-1998; 98WO-IB001665.
XX
XX 06-NOV-1997; 97GB-00023516.
PR 14-NOV-1997; 97GB-00024190.
PR 18-NOV-1997; 97GB-00024386.
PR 27-NOV-1997; 97GB-00025158.
PR 10-DEC-1997; 97GB-00026147.
PR 14-JAN-1998; 98GB-00000759.
PR 01-SEP-1998; 98GB-00019016.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
PI
XX WPI; 1999-327407/27.
XX P-PSDB; AAY38591.
XX
XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection.
XX
XX Claim 9; Page 150-151; 524pp; English.
XX
XX Nucleotide sequences AAZ11972-Z12358 represent open reading frames (ORFs)
CC of Neisseria meningitidis and N. gonorrhoeae which encode antigenic
CC proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,
CC their nucleic acids and antibodies are used for diagnosis, prevention (as
CC vaccines) or treatment of Neisseria infections, such as meningitis,
CC septicaemia and gonorrhea. Both organisms are closely related. Fragments
CC of the nucleic acids are useful as hybridisation probes and antisense
CC reagents
XX
XX Sequence 792 BP; 210 A; 184 C; 189 G; 209 T; 0 U; 0 Other;
SQ
Query Match 84.0%; Score 16.8; DB 2; Length 792;
Best Local Similarity 90.0%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAGGATAGAAACG 20
DB 148 GAGTAGGAGGATAGAAACG 129
|||||
|||||

RESULT 7
ID AEB49028 standard; DNA; 792 BP.
AC AEB49028;
XX
XX 22-SEP-2005 (first entry)
DT
XX
XX N. gonorrhoeae DNA SEQ ID NO 185.
DE
XX
XX antibacterial; immunosuppressive; antiinflammatory; vaccine;
KW neisseria gonorrhoeae infection; gynecological; infection;
KW neisseria meningitidis infection; meningitis; neuroprotective;
KW inflammation; neurological disease; sepsis; diagnosis; ds; gene.
XX
XX Neisseria gonorrhoeae.
OS
XX
XX US6914131-B1.
XX
XX 05-JUL-2005.
PD
XX
XX 30-APR-1999; 99US-00303518.
PF
XX
XX 09-OCT-1998; 98WO-IB001665.
PR
XX (CHIR ) CHIRON SRL.
XX
XX Scarlato V, Masignani V, Rappuoli R, Pizza M, Grandi G;
PI
XX WPI; 2005-464877/47.
XX
XX P-PSDB; AEB49029.
DR

```

```

XX
XX New Neisserial nucleic acids useful for diagnosing and/or treating
PT bacterial infections, in particular meningitis and septicaemia caused by
PT Neisseria meningitidis and Neisseria gonorrhoea.
XX
XX Example 22; SEQ ID NO 185; 613pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule. The methods
CC and compositions of the present invention are useful for diagnosing
CC and/or treating Neisserial bacterial infections, in particular meningitis
CC and septicaemia caused by Neisseria meningitidis and Neisseria gonorrhoea.
CC The present sequence represents a N. gonorrhoeae DNA. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=6914131B1.
XX
XX Sequence 792 BP; 210 A; 184 C; 189 G; 209 T; 0 U; 0 Other;
SQ
Query Match 84.0%; Score 16.8; DB 14; Length 792;
Best Local Similarity 90.0%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAGGATAGAAACG 20
DB 148 GAGTAGGAGGATAGAAACG 129
|||||
|||||

RESULT 8
ID ACA41244/c
XX ACA41244 standard; DNA; 2086 BP.
XX
XX ACA41244;
XX
XX 27-OCT-2003 (revised)
DT 19-JUN-2003 (first entry)
DT
XX
XX Prokaryotic essential gene #22901.
DE
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
KW
XX
XX Neisseria gonorrhoeae.
OS
XX
XX WO200277183-A2.
PN
XX
XX 03-OCT-2002.
PD
XX
XX 21-MAR-2002; 2002WO-US009107.
PF
XX
XX 21-MAR-2001; 2001US-00815242.
PR
XX 06-SEP-2001; 2001US-00948993.
PR
XX 25-OCT-2001; 2001US-0342923P.
PR
XX 08-FEB-2002; 2002US-00072851.
PR
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX
XX P-PSDB; ABU37374.
DR
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 29114; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC

```

CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)

XX SQ Sequence 2086 BP; 643 A; 461 C; 493 G; 489 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 8; Length 2086;
Best Local Similarity 90.0%; Pred. No. 6.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20
|||||
Db 148 GAGTAGGAATGAAAGAAACG 129

RESULT 9
AAZ53858/c
ID AAZ53858 standard; DNA; 6147 BP.
XX AC AAZ53858;
XX 15-SEP-2003 (revised)
DT 21-MAR-2000 (first entry)
XX Neisseria gonorrhoeae ORF 564 partial DNA sequence SEQ ID NO:1665.
DE Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX Neisseria gonorrhoeae.
OS WO9957280-A2.
XX PN 11-NOV-1999.
XX 30-APR-1999; 99WO-0009346.
XX 01-MAY-1998; 98US-0083758P.
PR 31-JUL-1998; 98US-0094869P.
PR 02-SEP-1998; 98US-0098994P.
PR 02-SEP-1998; 98US-0098062P.
PR 09-OCT-1998; 98US-0103749P.
PR 09-OCT-1998; 98US-0103794P.
PR 09-OCT-1998; 98US-0103796P.
PR 25-FEB-1999; 99US-0121528P.
PA (CHIR) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX Tettelin H, Venter JC;
DR WPI; 2000-062150/05.
DR P-PSDB; AAY75096.
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics.
PS Claim 7; Page 849-851; 1453pp; English.
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to Neisserial
CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC Neisseria bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols. (Updated on 15-SEP-2003 to standardise OS
CC field)

XX SQ Sequence 6147 BP; 1998 A; 1476 C; 1444 G; 1229 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 3; Length 6147;
Best Local Similarity 90.0%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20
|||||
Db 148 GAGTAGGAATGAAAGAAACG 129

RESULT 10
AAK84974/c
ID AAK84974 standard; DNA; 35871 BP.
XX AC AAK84974;
XX 07-NOV-2001 (first entry)
DT Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39786.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytosolic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS WO200157182-A2.
XX PN 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US001354.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184684P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234978P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 03-DEC-2000; 2000US-0251030P.
PR 03-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-483426/52.
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and metastasis.
Disclosure; SEQ ID NO 39786; 307lpp + Sequence Listing; English.
AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
activity, and can be used in gene therapy and vaccine production. (I)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (I) expression. For
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of (I) by expressing inactive proteins or to
supplement the patients own production of (I). Additionally, (I)
polynucleotides may be used to produce the secreted (I), by inserting the

CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above). The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterising
 CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 CC
 XX
 SQ Sequence 102 BP; 27 A; 15 C; 22 G; 38 T; 0 U; 0 Other;
 Query Match 82.0%; Score 16.4; DB 12; Length 102;
 Best Local Similarity 94.4%; Pred. No. 8.1e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GAGTAGGAGGATAGAAA 18
 DB 85 GAGAAGGAGGATAGAAA 68
 ||| ||||| ||||| |||||
 AC ABN23941;
 XX
 DT 24-JUN-2002 (first entry)
 XX
 DE Human ORFX polynucleotide sequence SEQ ID NO:16359.
 XX
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis; gene; ss.
 KW
 OS Homo sapiens.
 XX
 XX WO200192523-A2.
 FN
 XX
 PD 06-DEC-2001.
 XX
 XX 29-MAY-2001; 2001WO-US010836.
 PF
 XX 30-MAY-2000; 2000US-0206132P.
 PR
 XX 29-AUG-2000; 2000US-0228716P.
 PR
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shinkets RA, Leach MD;
 PI
 XX WPI; 2002-106308/14.
 DR
 XX P-PSDB; ABP08189.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.
 XX
 PS Disclosure; SEQ ID NO 16359; 1037pp; English.
 XX
 CC

CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, rheumatoid
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 XX
 SQ Sequence 316 BP; 138 A; 48 C; 64 G; 66 T; 0 U; 0 Other;
 Query Match 82.0%; Score 16.4; DB 6; Length 316;
 Best Local Similarity 94.4%; Pred. No. 8.6e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGTAGGAGGATAGAAC 19
 DB 184 AGTAGGAGGAGGAGAAC 201
 ||||| ||||| ||||| |||||
 AC AAC19786;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 23861.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 KW
 OS Homo sapiens.
 XX
 XX EF1033401-A2.
 FN
 XX
 PD 06-SEP-2000.
 XX
 XX 21-FEB-2000; 2000EP-00200610.
 PF
 XX 26-FEB-1999; 99US-0122487P.
 PR
 XX (GEST) GENSET.
 PA
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 PI
 XX WPI; 2000-500381/45.
 DR
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 1; SEQ ID NO 23861; 71pp + Sequence Listing; English.
 XX
 XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC

CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 10 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
 CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
 CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors

CC Sequence 347 BP; 132 A; 51 C; 57 G; 105 T; 0 U; 2 Other;

Query Match 82.0%; Score 16.4; DB 3; Length 347;
 Best Local Similarity 94.4%; Pred. No. 8.6e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAGGATAGAAAC 19

Db 246 AGTAGGAGGACAGAAAC 263

RESULT 15

ACH80283/C
 ID ACH80283 standard; DNA; 540 BP.

XX ACH80283;

XX 29-JUL-2004 (first entry)

XX Human genome derived single exon probe #13478.

XX Human; probe; ss; gene expression; single exon probe; microarray;
 XX alternative splicing event; genomic alteration.

XX Homo sapiens.

XX US2003194704-A1.

XX 16-OCT-2003.

XX 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.

XX (RANK/) RANK D R.

XX (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;

XX WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human
 XX gene expression analysis, for identifying or characterizing alternative
 XX splicing events, for assessing genomic alterations or as tools for
 XX surveying tissues.

XX Claim 15; SEQ ID NO 13478; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene
 XX expression, comprising any of the 27,400 fully defined nucleotide
 XX sequences in the specification, or their complements or fragments, and
 XX encoding at least 8 amino acids of any of the 688 amino acid sequences
 XX fully defined in the specification. The probe is a single exon probe that
 XX hybridises under high stringency conditions to a nucleic acid molecule
 XX expressed in human cells or tissues. Also included are a spatially-
 XX addressable set of single exon nucleic acid probes for measuring human
 XX gene expression (comprising a plurality of single exon nucleic acid
 XX probes cited above, where each of the plurality of probes is separately
 XX and addressably isolatable or amplifiable from the plurality), a single

CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterising
 CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX SQ Sequence 540 BP; 170 A; 110 C; 79 G; 181 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 12; Length 540;

Best Local Similarity 94.4%; Pred. No. 8.8e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATAGAAA 18

Db 217 GAGGAGGAGGATAGAAA 200

Search completed: January 2, 2006, 14:05:34

Job time : 312 secs

SHIT
GAVE
SIS
BUT
ALWAYS
KNOW

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 13:57:24 ; Search time 2243 Seconds
(without alignments)
417.183 Million cell updates/sec

Title: US-09-675-650-4

Perfect score: 20

Sequence: 1 gagtgaagatagaacg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gse1:*
10: gb_gse2:*
11: gb_gse3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	19	95.0	839	10	CZ978973 195767 To
c 2	18.4	92.0	690	10	CL161954 104 353 1
c 3	18.4	92.0	706	10	CW385789 fab5001f0
c 4	18.4	92.0	776	10	CL158398 104 347 1
c 5	18.4	92.0	812	10	CW385790 fab5001f0
c 6	17.4	87.0	230	5	BW856680 BW856680
c 7	17.4	87.0	278	5	BW855726 BW855726
c 8	17.4	87.0	475	5	BW853561 BW853561
c 9	17.4	87.0	488	5	BW850082 BW850082
c 10	17.4	87.0	517	5	BW874566 BW874566
c 11	17.4	87.0	525	5	BW871732 BW871732
c 12	17.4	87.0	533	2	BG49498 1024025E1
c 13	17.4	87.0	551	5	BW741437 BW741437
c 14	17.4	87.0	551	5	BW881148 BW881148
c 15	17.4	87.0	604	5	BW739479 BW739479
c 16	17.4	87.0	616	5	BW320697 BW320697
c 17	17.4	87.0	658	9	CC104842 CSU-K34.1
c 18	17.4	87.0	941	10	CL076406 CH216-139
c 19	17	85.0	549	10	C2439027 OA BBA009
c 20	16.8	84.0	180	10	BX650144 Arabidops
c 21	16.8	84.0	251	6	CA778813 MPL384.11
c 22	16.8	84.0	341	5	BQ846596 QGA19018.

23	16.8	84.0	343	5	BQ860812 QGC16M01.
24	16.8	84.0	361	5	BU007809 QGH5118.Y
25	16.8	84.0	362	5	BQ845348 QGA16H17
26	16.8	84.0	362	5	BQ847894 QGA5K10.Y
27	16.8	84.0	363	5	BQ859287 QGC12024.
28	16.8	84.0	365	5	BQ85414 QGS6M01.Y
29	16.8	84.0	367	5	BU003502 QGG35H06.
30	16.8	84.0	369	5	BQ985048 QGS5M10.Y
31	16.8	84.0	370	5	BQ856409 QGB3G11.Y
32	16.8	84.0	371	5	BQ981159 QGB13A05.
33	16.8	84.0	373	5	BQ853351 QGB20E10.
34	16.8	84.0	373	5	BQ862087 QGC20B21.
35	16.8	84.0	375	5	BQ847845 QGA5113.Y
36	16.8	84.0	376	5	BQ984885 QGS5F13.Y
37	16.8	84.0	405	1	AW982117 SSS0027.S
c 38	16.8	84.0	411	2	BE859183 SSS0490.S
c 39	16.8	84.0	427	5	BQ988279 QGF14H22.
c 40	16.8	84.0	439	5	BY270225 BY270225
c 41	16.8	84.0	511	2	BE810166 SSS0388.S
c 42	16.8	84.0	519	5	BY257655 BY257655
c 43	16.8	84.0	529	5	BQ861140 QGC17K08.
c 44	16.8	84.0	536	5	BQ983772 QGB1B03.Y
45	16.8	84.0	550	5	BQ852727 QGB18N05.

ALIGNMENTS

RESULT 1

CZ978973/c

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CZ978973 839 bp DNA linear GSS 11-AUG-2005
195767 Tomato MboI BAC Library Lycopersicon esculentum genomic
Clone SL_MboI0062A23 3, genomic survey sequence.

CZ978973

GSS.

LYCOPERSICON ESCULENTUM (Solanum lycopersicum)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 839)

1 (bases 1 to 839)

Van Eck, J., Buel, R.M., Wang, Y., Tanksley, S.D., Giovannoni, J.J.,

BAC end sequencing from three Solanum lycopersicon libraries

Unpublished (2005)

Other GSSs: 180444

Contact: Lukas Mueller

Tanksley Lab, Dept. of Plant Breeding

Cornell University

251 Emerson Hall, Ithaca, NY 14853, USA

Tel: 607-255-6557

Fax: 607-255-6683

Email: segn-feedback@sgn.cornell.edu

Plate: 62 row: A column: 23

Seq primer: SP6

Class: BAC ends

High quality sequence start: 106

High quality sequence stop: 359.

Location/Qualifiers

1. 839

/organism="Lycopersicon esculentum"

/mol_type="genomic DNA"

/cultivar="Heinz 1706"

/db_xref="taxon:4081"

/clone="SL_MboI0062A23"

/lab_host="E. coli"

/clone.lib="Tomato MboI BAC Library"

/note="Vector: pBelBAC11; Site_1: MboI"

ORIGIN

Query Match

Best Local Similarity

95.0%; Score 19; DB 10; Length 839;

100.0%; Pred. No. 3.3e+02;

```

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAAC 19
    |||||
Db 813 GAGTAGGAAGGATAGAAAC 795
    |||||

RESULT 2
CL161954/c
LOCUS 104_353.10806081.114 31828 369 Sorghum methylation-filtered library
DEFINITION (LibID: 104) Sorghum bicolor genomic clone 10806081, genomic survey
sequence.
ACCESSION CL161954
VERSION CL161954.1 GI:40668042
KEYWORDS GSS.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 690)
AUTHORS Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
McMenamy,J., Smith,M., Holean,H., Roe,B.A., Wiley,G., Korf,I.F.,
Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
Martienssen,R.A.
TITLE Sorghum genome sequencing by methylation filtration
JOURNAL PLoS Biol. 3 (1), e13 (2005)
PUBMED 15660154
COMMENT Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 353 row: p column: 09
Seq primer: M13/pUC Forward
Class: methylation filtered
High quality sequence stop: 690.
FEATURES
    source
    Location/Qualifiers
        1..690
            /organism="Sorghum bicolor"
            /mol_type="genomic DNA"
            /cultivar="ATx623"
            /db_xref="taxon:4558"
            /clone_lib="Sorghum methylation-filtered library (LibID:
            104)"
            /note="Organ: leaf; Vector: pBSCSK(-); Site 1: HincII; DNA
            prepared from purified nuclei was randomly sheared,
            end-repaired, size fractionated to enrich for the 0.5 to 5
            kb fraction, ligated into HincII-digested pBSCSK(-) vector
            and electroporated into E. coli cells. This is a
            methylation-filtered library."

ORIGIN
Query Match 92.0%; Score 18.4; DB 10; Length 690;
Best Local Similarity 95.0%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAACG 20
    |||||
Db 130 GAGCAGGAGGATAGAAACG 111
    |||||

RESULT 3
CW385789
LOCUS 706 bp DNA linear GSS 01-NOV-2004
DEFINITION fbb001f06909f0 Sorghum methylation filtered library (libid: 104)
Sorghum bicolor genomic clone fbb001f06909, genomic survey
sequence.
ACCESSION CW385789

```

```

CW385789.1 GI:55104233
GSS.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 706)
AUTHORS Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
McMenamy,J., Smith,M., Holean,H., Roe,B.A., Wiley,G., Korf,I.F.,
Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
Martienssen,R.A.
TITLE Sorghum genome sequencing by methylation filtration
JOURNAL PLoS Biol. 3 (1), e13 (2005)
PUBMED 15660154
COMMENT Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fbb001f069 row: g column: 09
Seq primer: f Forward
Class: methylation filtered
High quality sequence stop: 706.
FEATURES
    source
    Location/Qualifiers
        1..706
            /organism="Sorghum bicolor"
            /mol_type="genomic DNA"
            /cultivar="ATx623"
            /db_xref="taxon:4558"
            /clone_lib="Sorghum methylation filtered library (LibID:
            104)"
            /note="Organ: leaf; Vector: pBSCSK(-); Site 1: HincII; DNA
            prepared from purified nuclei was randomly sheared,
            end-repaired, size fractionated to enrich for the 0.5 to 5
            kb fraction, ligated into HincII-digested pBSCSK(-) vector
            and electroporated into E. coli cells. This is a
            methylation filtered library."

ORIGIN
Query Match 92.0%; Score 18.4; DB 10; Length 706;
Best Local Similarity 95.0%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAACG 20
    |||||
Db 563 GAGCAGGAGGATAGAAACG 582
    |||||

RESULT 4
CL158398/c
LOCUS 776 bp DNA linear GSS 06-JAN-2004
DEFINITION 104_347.10803514.114 31377 106 Sorghum methylation-filtered library
(LibID: 104) Sorghum bicolor genomic clone 10803514, genomic survey
sequence.
ACCESSION CL158398
VERSION CL158398.1 GI:40660972
KEYWORDS GSS.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 776)
AUTHORS Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
McMenamy,J., Smith,M., Holean,H., Roe,B.A., Wiley,G., Korf,I.F.,
Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
Martienssen,R.A.
TITLE Sorghum genome sequencing by methylation filtration

```

JOURNAL
PUBMED
COMMENT

PLoS Biol. 3 (1), e13 (2005)
15660154
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 347 row; e column: 10
Seq primer: M13/pUC Forward
Class: methylation filtered
High quality sequence stop: 776.

FEATURES

source

1. 776
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clones="10803514"
/clone_lib="Sorghum methylation-filtered library (LibID: 104)"
/note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation-filtered library."

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 776;
Best Local Similarity 95.0%; Pred. No. 6.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20

|||||

Db 425 GAGCAGGAAGGATAGAAACG 406

RESULT 5

CW385790/c

LOCUS

DEFINITION CW385790 812 bp DNA linear GSS 01-NOV-2004
fbb001f069g09k0 Sorghum methylation filtered library (LibID: 104)
Sorghum bicolor genomic clone fbb001f069g09, genomic survey
sequence.

ACCESSION CW385790

VERSION CW385790.1 GI:55104234

KEYWORDS

SOURCE

ORGANISM

Sorghum bicolor
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE

AUTHORS

1 (bases 1 to 812)
Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D.,
Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K.,
McKenamy, J., Smith, M., Holsman, H., Roe, B.A., Wiley, G., Korf, I.F.,
Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddle, J.A. and
Martienssen, R.A.

Sorghum genome sequencing by methylation filtration

PLoS Biol. 3 (1), e13 (2005)

JOURNAL

PUBMED

COMMENT

Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fbb001f069 row: 9 column: 09
Seq primer: k Reverse
Class: methylation filtered
High quality sequence stop: 812.

FEATURES

source

1. 812

/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clone_lib="Sorghum methylation filtered library (LibID: 104)"
/note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 812;
Best Local Similarity 95.0%; Pred. No. 6.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20

|||||

Db 625 GAGCAGGAAGGATAGAAACG 606

RESULT 6

BW856680

LOCUS

DEFINITION BW856680 Amphioxus Branchiostoma floridae unpublished cDNA library,
neurula whole animal Branchiostoma floridae cDNA clone bbne051a19
5', mRNA sequence.

ACCESSION BW856680

VERSION BW856680.1 GI:66464896

KEYWORDS

SOURCE

ORGANISM

Branchiostoma floridae (Florida lancelet)
Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 230)
Yu, J., Holland, J.Z., Shin-i, T., Kohara, Y., Satou, Y. and Satoh, N.
Expressed genes in Branchiostoma floridae
Unpublished (2005)
Contact: Tadasi Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source

1. 230
/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clones="bbne051a19"
/dev stage="neurula"
/dev stage="neurula"
/clone_lib="Amphioxus Branchiostoma floridae unpublished
cDNA library, neurula whole animal"

ORIGIN

Query Match 87.0%; Score 17.4; DB 5; Length 230;
Best Local Similarity 94.7%; Pred. No. 1.5e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAACG 20

|||||

Db 87 AGTAGGAAGGATAGAAACG 105

RESULT 7

BW855726

LOCUS

DEFINITION BW855726 Amphioxus Branchiostoma floridae unpublished cDNA library,
neurula whole animal

neurula whole animal Branchiostoma floridae cDNA clone bbne048g21
5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BW855726
BW855726.1 GI:66463942
EST.
Branchiostoma floridae (Florida lancelet)
Branchiostoma floridae

Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.

REFERENCE

1 (bases 1 to 278)
Yu, J., Holland, L.Z., Shin-i, T., Kohara, Y., Satou, Y. and Satoh, N.
Expressed genes in Branchiostoma floridae

Unpublished (2005)

COMMENT

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. .278

/organism="Branchiostoma floridae"

/mol_type="mRNA"

/db_xref="taxon:7739"

/clone="bbne048g21"

/tissue_type="whole animal"

/dev_stage="neurula"

/clone_lib="Amphioxus Branchiostoma floridae unpublished

cDNA library, neurula whole animal"

ORIGIN

Query Match 87.0%; Score 17.4; DB 5; Length 278;

Best Local Similarity 94.7%; Pred. No. 1.6e+03;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAACG 20

|||||

Db 11 AGTAGGAAGTATAGAAACG 29

RESULT 8

BW853561

LOCUS

DEFINITION

BW853561 Amphioxus Branchiostoma floridae (Florida lancelet)

neurula whole animal Branchiostoma floridae unpublished cDNA library,

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;

Branchiostoma.

REFERENCE

1 (bases 1 to 475)

Yu, J., Holland, L.Z., Shin-i, T., Kohara, Y., Satou, Y. and Satoh, N.

Expressed genes in Branchiostoma floridae

Unpublished (2005)

COMMENT

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. .475

/organism="Branchiostoma floridae"

/mol_type="mRNA"

/db_xref="taxon:7739"

/clone="bbne042f21"

/tissue_type="whole animal"

/dev_stage="neurula"

FEATURES

source

ORIGIN

Query Match

Best Local Similarity 87.0%; Score 17.4; DB 5; Length 475;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAACG 20

|||||

Db 27 AGTAGGAAGTATAGAAACG 45

RESULT 9

BW850082

LOCUS

DEFINITION

BW850082 Amphioxus Branchiostoma floridae (Florida lancelet)

neurula whole animal Branchiostoma floridae unpublished cDNA library,

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;

Branchiostoma.

REFERENCE

1 (bases 1 to 488)

Yu, J., Holland, L.Z., Shin-i, T., Kohara, Y., Satou, Y. and Satoh, N.

Expressed genes in Branchiostoma floridae

Unpublished (2005)

JOURNAL

COMMENT

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. .488

/organism="Branchiostoma floridae"

/mol_type="mRNA"

/db_xref="taxon:7739"

/clone="bbne038124"

/tissue_type="whole animal"

/dev_stage="neurula"

/clone_lib="Amphioxus Branchiostoma floridae unpublished

cDNA library, neurula whole animal"

ORIGIN

Query Match

Best Local Similarity 87.0%; Score 17.4; DB 5; Length 488;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAACG 20

|||||

Db 349 AGTAGGAAGTATAGAAACG 367

RESULT 10

BW874566

LOCUS

DEFINITION

BW874566 Amphioxus Branchiostoma floridae (Florida lancelet)

neurula whole animal Branchiostoma floridae unpublished cDNA library,

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;

Branchiostoma.

REFERENCE

1 (bases 1 to 517)

Yu, J., Holland, L.Z., Shin-i, T., Kohara, Y., Satou, Y. and Satoh, N.

```

TITLE      Expressed genes in Branchiostoma floridae
JOURNAL    Unpublished (2005)
COMMENT    Contact: Tadasu Shin-i
           National Institute of Genetics
           1111 Yata, Mishima, Shizuoka 411-8540, Japan
           Tel: 81-559-81-6856
           Fax: 81-559-81-6855
           Email: tshini@genes.nig.ac.jp.

FEATURES   source
            Location/Qualifiers
            1..517
            /organism="Branchiostoma floridae"
            /mol_type="mRNA"
            /db_xref="taxon:7739"
            /clone="bbnel1491"
            /tissue_type="whole animal"
            /dev_stage="neurula"
            /clone_lib="Amphioxus Branchiostoma floridae unpublished
            cDNA library, neurula whole animal"

ORIGIN
Query Match      87.0%; Score 17.4; DB 5; Length 517;
Best Local Similarity 94.7%; Pred. NO. 1.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2 AGTAGGAGGATAGAACG 20
    |||||
Db  317 AGTAGGAGGATAGAACG 335

RESULT 11
BW871732
LOCUS      525 bp mRNA linear EST 24-MAY-2005
DEFINITION Branchiostoma floridae (Florida lancelet)
            neurula whole animal Branchiostoma floridae cDNA clone bbnel106c12
            5', mRNA sequence.
ACCESSION  BW871732
VERSION     BW871732.1 GI:66486409
KEYWORDS   EST.
SOURCE     Branchiostoma floridae (Florida lancelet)
ORGANISM   Branchiostoma floridae
            Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
            Branchiostoma.
REFERENCE   1 (bases 1 to 525)
AUTHORS    Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE      Expressed genes in Branchiostoma floridae
JOURNAL    Unpublished (2005)
COMMENT    Contact: Tadasu Shin-i
           Center For Genetic Resource Information
           National Institute of Genetics
           1111 Yata, Mishima, Shizuoka 411-8540, Japan
           Tel: 81-559-81-6856
           Fax: 81-559-81-6855
           Email: tshini@genes.nig.ac.jp.

FEATURES   source
            Location/Qualifiers
            1..525
            /organism="Branchiostoma floridae"
            /mol_type="mRNA"
            /db_xref="taxon:7739"
            /clone="bbnel106c12"
            /tissue_type="whole animal"
            /dev_stage="neurula"
            /clone_lib="Amphioxus Branchiostoma floridae unpublished
            cDNA library, neurula whole animal"

ORIGIN
Query Match      87.0%; Score 17.4; DB 5; Length 525;
Best Local Similarity 94.7%; Pred. NO. 1.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2 AGTAGGAGGATAGAACG 20
    |||||
Db  228 AGTAGGAGGATAGAACG 246

```

```

RESULT 12
BG849498/c
LOCUS      533 bp mRNA linear EST 29-MAY-2001
DEFINITION Chlamydomonas reinhardtii CC-1690, normalized, Lambda Zap II
            Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION  BG849498
VERSION     BG849498.1 GI:14230682
KEYWORDS   EST.
SOURCE     Chlamydomonas reinhardtii
            Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
            Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1 (bases 1 to 533)
AUTHORS    Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
            McDermott,J.P., Sllflow,C., Stern,D. and Surzycki,R.
TITLE      Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants; project phase 2
JOURNAL    Unpublished (2000)
COMMENT    Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.

FEATURES   source
            Location/Qualifiers
            1..533
            /organism="Chlamydomonas reinhardtii"
            /mol_type="mRNA"
            /strain="CC-1690 wild type mt+ 21gr"
            /db_xref="taxon:3055"
            /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
            II"
            /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
            XhoI; This library, constructed by John Davies and Jeffrey
            McDermott, combines cDNAs from CC-1690 cells grown to
            mid-log phase in TAP (acetate-containing) medium in the
            light, TAP medium in the dark, HS (minimal) medium in
            ambient levels of CO2 and HS medium bubbled with 5% CO2.
            polyA mRNA was purified from each sample, pooled and cDNA
            synthesized. The cDNA was directionally cloned into lambda
            ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
            pBluescript II SK- plasmids were excised from the lambda
            ZAP clones by superinfection with Exassist (Stratagene)
            phage. The library was normalized using method 4 described
            in Bonaldo et al (1996) Genome Research 6: 791-806."

ORIGIN
Query Match      87.0%; Score 17.4; DB 2; Length 533;
Best Local Similarity 94.7%; Pred. NO. 1.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1 GACTAGGAGGATAGAAC 19
    |||||
Db  321 GACTAGGAGGATAGAAC 303

RESULT 13
BW741437
LOCUS      551 bp mRNA linear EST 09-AUG-2005
DEFINITION Amphioxus Branchiostoma floridae unpublished cDNA library,
            egg whole animal Branchiostoma floridae cDNA clone bbeg039n13 5',
            mRNA sequence.
ACCESSION  BW741437
VERSION     BW741437.1 GI:66328085
KEYWORDS   EST.
SOURCE     Branchiostoma floridae (Florida lancelet)
            Branchiostoma floridae
            Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
            Branchiostoma.

```

```

REFERENCE 1 (bases 1 to 551)
AUTHORS Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE Expressed genes in Branchiostoma floridae
JOURNAL Unpublished (2005)
COMMENT Contact: Tadasu Shin-i
        Center For Genetic Resource Information
        National Institute of Genetics
        1111 Yata, Mishima, Shizuoka 411-8540, Japan
        Tel: 81-559-81-6856
        Fax: 81-559-81-6855
        Email: tshini@genes.nig.ac.jp
If you want to have a cDNA clone for this EST or if you have any
questions, please send an e-mail to Nori Satoh
(satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou
(yutaka@ascidian.zool.kyoto-u.ac.jp).
FEATURES             source
    source
    1..551
        /organism="Branchiostoma floridae"
        /mol_type="mRNA"
        /db_xref="taxon:7739"
        /clone="bbeg039n13"
        /tissue_type="whole animal"
        /dev_stage="egg"
        /clone_lib="Amphioxus Branchiostoma floridae unpublished
        cDNA library, egg whole animal"
ORIGIN
Query Match      87.0%; Score 17.4; DB 5; Length 551;
Best Local Similarity 94.7%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAACG 20
|||||
Db 97 AGTAGGAAGGATAGAAACG 115

RESULT 14
BW881148 551 bp mRNA linear EST 24-MAY-2005
LOCUS BW881148
DEFINITION BW881148 Amphioxus Branchiostoma floridae unpublished cDNA library,
neurula whole animal Branchiostoma floridae cDNA clone bbne095n04
5', mRNA sequence.
ACCESSION BW881148
VERSION BW881148.1 GI:66495825
KEYWORDS EST.
SOURCE Branchiostoma floridae (Florida lancelet)
ORGANISM Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
REFERENCE 1 (bases 1 to 551)
AUTHORS Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE Expressed genes in Branchiostoma floridae
JOURNAL Unpublished (2005)
COMMENT Contact: Tadasu Shin-i
        Center For Genetic Resource Information
        National Institute of Genetics
        1111 Yata, Mishima, Shizuoka 411-8540, Japan
        Tel: 81-559-81-6856
        Fax: 81-559-81-6855
        Email: tshini@genes.nig.ac.jp.
FEATURES             source
    source
    1..551
        /organism="Branchiostoma floridae"
        /mol_type="mRNA"
        /db_xref="taxon:7739"
        /clone="bbne095n04"
        /tissue_type="whole animal"
        /dev_stage="neurula"
        /clone_lib="Amphioxus Branchiostoma floridae unpublished
        cDNA library, neurula whole animal"
ORIGIN
Query Match      87.0%; Score 17.4; DB 5; Length 551;
Best Local Similarity 94.7%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAACG 20
|||||
Db 97 AGTAGGAAGGATAGAAACG 115

RESULT 14
BW881148 551 bp mRNA linear EST 24-MAY-2005
LOCUS BW881148
DEFINITION BW881148 Amphioxus Branchiostoma floridae unpublished cDNA library,
neurula whole animal Branchiostoma floridae cDNA clone bbne095n04
5', mRNA sequence.
ACCESSION BW881148
VERSION BW881148.1 GI:66495825
KEYWORDS EST.
SOURCE Branchiostoma floridae (Florida lancelet)
ORGANISM Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
REFERENCE 1 (bases 1 to 551)
AUTHORS Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE Expressed genes in Branchiostoma floridae
JOURNAL Unpublished (2005)
COMMENT Contact: Tadasu Shin-i
        Center For Genetic Resource Information
        National Institute of Genetics
        1111 Yata, Mishima, Shizuoka 411-8540, Japan
        Tel: 81-559-81-6856
        Fax: 81-559-81-6855
        Email: tshini@genes.nig.ac.jp.
FEATURES             source
    source
    1..604
        /organism="Branchiostoma floridae"
        /mol_type="mRNA"
        /db_xref="taxon:7739"
        /clone="bbeg025f07"
        /tissue_type="whole animal"
        /dev_stage="egg"
        /clone_lib="Amphioxus Branchiostoma floridae unpublished
        cDNA library, egg whole animal"
ORIGIN
Query Match      87.0%; Score 17.4; DB 5; Length 604;
Best Local Similarity 94.7%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAACG 20
|||||
Db 161 AGTAGGAAGGATAGAAACG 179

Search completed: January 2, 2006, 15:18:40
Job time : 2246 secs

```

```

Best Local Similarity 94.7%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAACG 20
|||||
Db 432 AGTAGGAAGGATAGAAACG 450

RESULT 15
BW739479 604 bp mRNA linear EST 09-AUG-2005
LOCUS BW739479
DEFINITION BW739479 Amphioxus Branchiostoma floridae unpublished cDNA library,
egg whole animal Branchiostoma floridae cDNA clone bbeg025f07 5',
mRNA sequence.
ACCESSION BW739479
VERSION BW739479.1 GI:66326109
KEYWORDS EST.
SOURCE Branchiostoma floridae (Florida lancelet)
ORGANISM Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
REFERENCE 1 (bases 1 to 604)
AUTHORS Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE Expressed genes in Branchiostoma floridae
JOURNAL Unpublished (2005)
COMMENT Contact: Tadasu Shin-i
        Center For Genetic Resource Information
        National Institute of Genetics
        1111 Yata, Mishima, Shizuoka 411-8540, Japan
        Tel: 81-559-81-6856
        Fax: 81-559-81-6855
        Email: tshini@genes.nig.ac.jp
If you want to have a cDNA clone for this EST or if you have any
questions, please send an e-mail to Nori Satoh
(satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou
(yutaka@ascidian.zool.kyoto-u.ac.jp).
FEATURES             source
    source
    1..604
        /organism="Branchiostoma floridae"
        /mol_type="mRNA"
        /db_xref="taxon:7739"
        /clone="bbeg025f07"
        /tissue_type="whole animal"
        /dev_stage="egg"
        /clone_lib="Amphioxus Branchiostoma floridae unpublished
        cDNA library, egg whole animal"
ORIGIN
Query Match      87.0%; Score 17.4; DB 5; Length 604;
Best Local Similarity 94.7%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAACG 20
|||||
Db 161 AGTAGGAAGGATAGAAACG 179

Search completed: January 2, 2006, 15:18:40
Job time : 2246 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 14:00:00 ; Search time 104 Seconds
(without alignments)
341.839 Million cell updates/sec

Title: US-09-675-650-4

Perfect score: 20

Sequence: 1 gaggtaggaagatagaacg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB seq:*
- 6: /cgn2_6/ptodata/1/ina/PP COMB seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.4	87.0	601	3	US-09-949-016-40733 Sequence 40733, A
C 2	17.4	87.0	462589	3	US-09-949-016-12900 Sequence 12900, A
C 3	17.4	87.0	476044	3	US-09-949-016-12412 Sequence 12412, A
4	16.4	82.0	347	3	US-09-513-999C-23861 Sequence 23861, A
5	16.4	82.0	312474	3	US-09-949-016-17434 Sequence 17434, A
6	16	80.0	117838	3	US-09-949-016-17595 Sequence 17595, A
C 7	15.8	79.0	601	3	US-09-949-016-117449 Sequence 117449, A
C 8	15.8	79.0	601	3	US-09-949-016-169788 Sequence 169788, A
9	15.8	79.0	41736	3	US-09-949-016-17091 Sequence 17091, A
10	15.8	79.0	61178	3	US-09-949-016-17369 Sequence 17369, A
11	15.8	79.0	67386	3	US-09-949-016-16519 Sequence 16519, A
C 12	15.8	79.0	94593	3	US-09-949-016-16324 Sequence 16324, A
C 13	15.8	79.0	111235	3	US-09-949-016-15328 Sequence 15328, A
C 14	15.8	79.0	114842	3	US-09-949-016-14993 Sequence 14993, A
15	15.4	77.0	601	3	US-09-949-016-92291 Sequence 92291, A
16	15.4	77.0	601	3	US-09-949-016-92292 Sequence 92292, A
17	15.4	77.0	601	3	US-09-949-016-92293 Sequence 92293, A
18	15.4	77.0	601	3	US-09-949-016-92294 Sequence 92294, A
C 19	15.4	77.0	601	3	US-09-949-016-11635 Sequence 11635, A
20	15.4	77.0	2973	3	US-10-104-047-634 Sequence 634, App
21	15.4	77.0	38009	3	US-09-949-016-13617 Sequence 13617, A
C 22	15.4	77.0	69737	3	US-09-949-016-15140 Sequence 15140, A
23	15.4	77.0	105413	3	US-10-427-923-3 Sequence 3, Appli
24	15.4	77.0	112219	3	US-09-949-016-12453 Sequence 12453, A

25	15.4	77.0	112222	3	US-09-949-016-14324 Sequence 14324, A
26	15.4	77.0	113186	3	US-09-949-016-17572 Sequence 17572, A
C 27	15.4	77.0	142504	3	US-09-949-016-13693 Sequence 13693, A
C 28	15.4	77.0	142506	3	US-09-949-016-12474 Sequence 12474, A
29	15.4	77.0	192956	3	US-09-949-016-14382 Sequence 14382, A
30	15.4	77.0	636591	3	US-09-949-016-11808 Sequence 11808, A
31	15.4	77.0	636591	3	US-09-949-016-13388 Sequence 13388, A
32	15.2	76.0	228	3	US-09-107-532A-1850 Sequence 1850, Ap
C 33	15.2	76.0	483	3	US-09-949-016-3457 Sequence 3457, Ap
C 34	15.2	76.0	570	3	US-09-107-532A-3368 Sequence 3368, Ap
C 35	15.2	76.0	601	3	US-09-949-016-26370 Sequence 26370, A
C 36	15.2	76.0	601	3	US-09-949-016-34217 Sequence 34217, A
C 37	15.2	76.0	601	3	US-09-949-016-34218 Sequence 34218, A
C 38	15.2	76.0	601	3	US-09-949-016-34219 Sequence 34219, A
C 39	15.2	76.0	601	3	US-09-949-016-34220 Sequence 34220, A
C 40	15.2	76.0	601	3	US-09-949-016-50910 Sequence 50910, A
C 41	15.2	76.0	601	3	US-09-949-016-50911 Sequence 50911, A
C 42	15.2	76.0	601	3	US-09-949-016-122574 Sequence 122574, A
C 43	15.2	76.0	601	3	US-09-949-016-122575 Sequence 122575, A
C 44	15.2	76.0	601	3	US-09-949-016-133874 Sequence 133874, A
C 45	15.2	76.0	601	3	US-09-949-016-133874 Sequence 133874, A

ALIGNMENTS

RESULT 1
US-09-949-016-40733
; Sequence 40733, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40733
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-40733

Query Match 87.0%; Score 17.4; DB 3; Length 601;
Best Local Similarity 94.7%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GAGTAGGAGGATAGAAAC 19
|||||
Db 446 GAGTAGGAGGAGGAGAAAC 464

RESULT 2
US-09-949-016-12900/c
; Sequence 12900, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12900
; LENGTH: 462589
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12900

Query Match 87.0%; Score 17.4; DB 3; Length 462589;
Best Local Similarity 94.7%; Pred. No. 93;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAAC 19
|||||
Db 132852 GAGTAGGAAGGAAAGAAAC 132834

RESULT 3

US-09-949-016-12412/c
; Sequence 12412, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12412
; LENGTH: 476044
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12412

Query Match 87.0%; Score 17.4; DB 3; Length 476044;
Best Local Similarity 94.7%; Pred. No. 93;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAAC 19
|||||
Db 146308 GAGTAGGAAGGAAAGAAAC 146290

RESULT 4

US-09-513-999C-23861
; Sequence 23861, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 23861

; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 311
; OTHER INFORMATION: r=a or g
; US-09-513-999C-23861

Query Match 82.0%; Score 16.4; DB 3; Length 347;
Best Local Similarity 94.4%; Pred. No. 11e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAAC 19
|||||
Db 246 AGTAGGAAGGACAGAAAC 263

RESULT 5

US-09-949-016-17434
; Sequence 17434, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17434
; LENGTH: 312474
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17434

Query Match 82.0%; Score 16.4; DB 3; Length 312474;
Best Local Similarity 94.4%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAAC 19
|||||
Db 9872 AGTTGAAGGATAGAAAC 9889

RESULT 6

US-09-949-016-17595
; Sequence 17595, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08


```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17595
; LENGTH: 117838
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(117838)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17595

Query Match      80.0%; Score 16; DB 3; Length 117838;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4 TAGGAAGGATAGAAAC 19
Db      15884 TAGGAAGGATAGAAAC 15899
|||||

RESULT 7
US-09-949-016-117449/c
; Sequence 117449, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117449
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-117449

Query Match      79.0%; Score 15.8; DB 3; Length 601;
Best Local Similarity 89.5%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GAGTAGGAAGGATAGAAAC 19
Db      498 GAGTAGGCTGGATAGAAAC 480
|||||

RESULT 8
US-09-949-016-169788/c
; Sequence 169788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169788
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-169788

Query Match      79.0%; Score 15.8; DB 3; Length 601;
Best Local Similarity 89.5%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GAGTAGGAAGGATAGAAAC 19
Db      2761 GAGTAGGAAGGATAGAGAC 2779
|||||

RESULT 9
US-09-949-016-17091
; Sequence 17091, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17091
; LENGTH: 41736
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17091

Query Match      79.0%; Score 15.8; DB 3; Length 41736;
Best Local Similarity 89.5%; Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GAGTAGGAAGGATAGAAAC 19
Db      2761 GAGTAGGAAGGATAGAGAC 2779
|||||

RESULT 10
US-09-949-016-17369
; Sequence 17369, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17369
; LENGTH: 61178
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17369
```

US-09-949-016-17369

Query Match 79.0%; Score 15.8; DB 3; Length 61178;
Best Local Similarity 89.5%; Pred. No. 4.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAAC 19
||||| ||||||| |||||
Db 44335 GAGTATGCGAGGATAGAAAC 44353

RESULT 11

US-09-949-016-16519
; Sequence 16519, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16519
; LENGTH: 67386
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(67386)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16519

Query Match 79.0%; Score 15.8; DB 3; Length 67386;
Best Local Similarity 89.5%; Pred. No. 4.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAAC 19
||||| ||||||| |||||
Db 3862 GACTAGGAAGGATAGAAAC 3880

RESULT 12

US-09-949-016-16324/c
; Sequence 16324, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16324
; LENGTH: 94593
; TYPE: DNA
; ORGANISM: Human
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: (1)---(94593)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16324

Query Match 79.0%; Score 15.8; DB 3; Length 94593;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAAC 19
||||| ||||||| |||||
Db 31140 GAGTAGGAAGGATAGAAAC 31122

RESULT 13

US-09-949-016-15328
; Sequence 15328, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15328
; LENGTH: 111235
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)---(111235)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15328

Query Match 79.0%; Score 15.8; DB 3; Length 111235;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGTAGGAAGGATAGAAACG 20
||||| ||||||| |||||
Db 23769 AGTAGGAAGGATAGAAACG 23787

RESULT 14

US-09-949-016-14993/c
; Sequence 14993, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14993
; LENGTH: 114842

```

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(114842)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14993

Query Match      79.0%; Score 15.8; DB 3; Length 114842;
Best Local Similarity 89.5%; Pred. No. 4.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GAGTAGGAAGGATAGAAAC 19
        |||||
Db      45824 GAGTAGGCTGGATAGAAAC 45806

RESULT 15
US-09-949-016-92291
; Sequence 92291, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92291
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-92291

Query Match      77.0%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 3.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GAGTAGGAAGGATAGAA 17
        |||||
Db      509 GAGTAGGAAGGATACAA 525

```

Search completed: January 2, 2006, 15:20:30
Job time : 107 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 14:05:41 ; Search time 451 Seconds
(without alignments)
366.713 Million cell updates/sec

Title: US-09-675-650-4

Perfect score: 20

Sequence: 1 gaggtagaagatagaacg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /cgm2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
- 2: /cgm2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
- 3: /cgm2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
- 4: /cgm2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
- 5: /cgm2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
- 6: /cgm2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
- 7: /cgm2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
- 8: /cgm2_6/ptodata/1/pubna/US10D_PUBCOMB.seq.*
- 9: /cgm2_6/ptodata/1/pubna/US10E_PUBCOMB.seq.*
- 10: /cgm2_6/ptodata/1/pubna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	17.4	87.0	348	7 US-10-424-599-12713	Sequence 12713, A
C 2	17.4	87.0	32323	5 US-10-087-192-376	Sequence 376, App
C 3	16.8	84.0	290	8 US-10-425-115-178767	Sequence 178767, A
C 4	16.8	84.0	553	4 US-09-925-065A-267785	Sequence 267785, A
C 5	16.8	84.0	2086	7 US-10-282-122A-29114	Sequence 29114, A
C 6	16.8	84.0	99924	6 US-10-085-117-115	Sequence 115, App
C 7	16.4	82.0	65	3 US-09-908-975-27906	Sequence 27906, A
C 8	16.4	82.0	102	6 US-10-029-386-27178	Sequence 27178, A
C 9	16.4	82.0	528	4 US-09-925-065A-542839	Sequence 542839, A
C 10	16.4	82.0	540	6 US-10-029-386-13478	Sequence 13478, A
C 11	16.4	82.0	549	4 US-09-925-065A-327859	Sequence 327859, A
C 12	16.4	82.0	582	4 US-09-925-065A-635703	Sequence 635703, A
C 13	16.4	82.0	600	9 US-10-972-079-67862	Sequence 67862, A
C 14	16.4	82.0	600	9 US-10-972-079-67863	Sequence 67863, A
C 15	16.4	82.0	770	4 US-09-925-065A-70545	Sequence 70545, A
C 16	16.4	82.0	2300	5 US-10-027-632-103227	Sequence 103227, A
C 17	16.4	82.0	2300	5 US-10-027-632-103228	Sequence 103228, A
C 18	16.4	82.0	2300	6 US-10-027-632-103227	Sequence 103227, A
C 19	16.4	82.0	2300	6 US-10-027-632-103228	Sequence 103228, A
C 20	16	80.0	417	7 US-10-424-599-133867	Sequence 133867, A
C 21	16	80.0	579	4 US-09-925-065A-543153	Sequence 543153, A
C 22	16	80.0	579	4 US-09-925-065A-543154	Sequence 543154, A
C 23	15.8	79.0	214	6 US-10-029-386-25656	Sequence 25656, A

C 24	15.8	79.0	241	7 US-10-424-599-123775	Sequence 123775, A
C 25	15.8	79.0	300	9 US-10-779-543-2283	Sequence 2283, App
C 26	15.8	79.0	304	8 US-10-357-930-59253	Sequence 59253, A
C 27	15.8	79.0	506	6 US-10-029-386-116222	Sequence 116222, A
C 28	15.8	79.0	512	4 US-09-925-065A-565386	Sequence 565386, A
C 29	15.8	79.0	515	7 US-10-767-701-29569	Sequence 29569, A
C 30	15.8	79.0	520	5 US-10-027-632-192849	Sequence 192849, A
C 31	15.8	79.0	520	5 US-10-027-632-192850	Sequence 192850, A
C 32	15.8	79.0	520	6 US-10-027-632-192849	Sequence 192849, A
C 33	15.8	79.0	520	6 US-10-027-632-192850	Sequence 192850, A
C 34	15.8	79.0	548	6 US-10-029-386-11936	Sequence 11936, A
C 35	15.8	79.0	569	4 US-09-925-065A-662453	Sequence 662453, A
C 36	15.8	79.0	569	4 US-09-925-065A-662454	Sequence 662454, A
C 37	15.8	79.0	569	4 US-09-925-065A-662455	Sequence 662455, A
C 38	15.8	79.0	569	4 US-09-925-065A-662456	Sequence 662456, A
C 39	15.8	79.0	579	7 US-10-425-114-36237	Sequence 36237, A
C 40	15.8	79.0	587	4 US-09-925-065A-253782	Sequence 253782, A
C 41	15.8	79.0	608	7 US-10-767-701-20245	Sequence 20245, A
C 42	15.8	79.0	642	4 US-09-925-065A-822921	Sequence 822921, A
C 43	15.8	79.0	643	4 US-09-925-065A-812843	Sequence 812843, A
C 44	15.8	79.0	969	5 US-10-027-632-31614	Sequence 31614, A
C 45	15.8	79.0	969	6 US-10-027-632-31614	Sequence 31614, A

ALIGNMENTS

RESULT 1
US-10-424-599-12713/c
; Sequence 12713, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12713
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(348)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_111487C.1
US-10-424-599-12713
Query Match 87.0%; Score 17.4; DB 7; Length 348;
Best Local Similarity 94.7%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGTAGGAGGATAGAAAC 19
|||||
Db 142 GAGTAGGAGGATAGAAAC 124
RESULT 2
US-10-087-192-376/c
; Sequence 376, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122

; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 376
; LENGTH: 32323
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(32323)
; OTHER INFORMATION: n = A,T,C or G
; US-10-087-192-376

Query Match 87.0%; Score 17.4; DB 5; Length 32323;
Best Local Similarity 94.7%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAAC 19
||||| ||||| ||||| ||||| |||||
Db 7698 GAGTAGGAAGGATAGAAAC 7680

RESULT 3

US-10-425-115-178767/c
; Sequence 178767, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 178767
; LENGTH: 290
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(290)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_94619C.1
; US-10-425-115-178767

Query Match 84.0%; Score 16.8; DB 8; Length 290;
Best Local Similarity 90.0%; Pred. No. 4.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAAC 20
||||| ||||| ||||| ||||| |||||
Db 58 GAGGAGGAAGGAAGAAAC 39

RESULT 4

US-09-925-065A-267785/c
; Sequence 267785, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A

; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 267785
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-267785

Query Match 84.0%; Score 16.8; DB 4; Length 553;
Best Local Similarity 90.0%; Pred. No. 4.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAAC 20
||||| ||||| ||||| ||||| |||||
Db 166 GAGGAGGAAGGATAGAAAC 147

RESULT 5

US-10-282-122A-29114/c
; Sequence 29114, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29114

```
; LENGTH: 2086
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-282-122A-29114

Query Match      84.0%; Score 16.8; DB 7; Length 2086;
Best Local Similarity 90.0%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20
    ||||| ||||| ||||| |||||
Db 148 GAGTAGGAATGAAAGAAACG 129

RESULT 6
US-10-085-117-115/c
; Sequence 115, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 99924
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(99924)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-115

Query Match      84.0%; Score 16.8; DB 6; Length 99924;
Best Local Similarity 90.0%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20
    ||||| ||||| ||||| |||||
Db 48494 GAGTAGGAAGGAGAGAAAGG 48475

RESULT 7
US-09-908-975-27906
; Sequence 27906, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27906
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
```

```
US-09-908-975-27906

Query Match      82.0%; Score 16.4; DB 3; Length 65;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAAC 19
    ||||| ||||| ||||| |||||
Db 19 AGTAGGAAGGATAGACAC 36

RESULT 8
US-10-029-386-27178/c
; Sequence 27178, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27178
; LENGTH: 102
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR14.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.63
; OTHER INFORMATION: SWISSPROT HIT: P20917, EVALUE 8.00e+00
; OTHER INFORMATION: NT HIT: AJ002571.1, EVALUE 1.30e+00
US-10-029-386-27178

Query Match      82.0%; Score 16.4; DB 6; Length 102;
Best Local Similarity 94.4%; Pred. No. 6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAA 18
    ||||| ||||| ||||| |||||
Db 85 GAGTAGGAAGGATAGAAA 68

RESULT 9
US-09-925-065A-542839
; Sequence 542839, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 542839
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-09-925-065A-542839
Query Match      82.0%; Score 16.4; DB 4; Length 528;
Best Local Similarity 94.4%; Pred. No. 6.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GTAGGAAGGATAGAAACG 20
   |||||
Db 101 GTAGGAAGGATAGAAATG 118

RESULT 10
US-10-029-386-13478/c
; Sequence 13478, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13478
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR14.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.63
; OTHER INFORMATION: NT HIT: U87252.1, EVALUAE 3.30e-02
; OTHER INFORMATION: SWISSPROT HIT: Q03560, EVALUAE 2.30e+00
US-10-029-386-13478

Query Match      82.0%; Score 16.4; DB 6; Length 540;
Best Local Similarity 94.4%; Pred. No. 6.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTAGGAAGGATAGAAA 18
   |||||
Db 217 GAGTAGGAAGGATAGAAA 200

RESULT 11
US-09-925-065A-327859/c
; Sequence 327859, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 327859
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-09-925-065A-327859
Query Match      82.0%; Score 16.4; DB 4; Length 549;
Best Local Similarity 94.4%; Pred. No. 6.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GTAGGAAGGATAGAAACG 20
   |||||
Db 285 GTAGGAAGGATAGAAACG 268

RESULT 12
US-09-925-065A-635703
; Sequence 635703, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 635703
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-635703

Query Match      82.0%; Score 16.4; DB 4; Length 582;
Best Local Similarity 94.4%; Pred. No. 6.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GTAGGAAGGATAGAAACG 20
   |||||
Db 544 GTAGGAAGGATAGAAATG 561

RESULT 13
US-10-972-079-67862/c
; Sequence 67862, Application US/10972079
; Publication No. US20050153317A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
; TITLE OF INVENTION: LIVESTOCK
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67862
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894338179_1
```


US-10-972-079-67862

Query Match 82.0%; Score 16.4; DB 9; Length 600;
Best Local Similarity 94.4%; Pred. No. 6.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAA 18
|||||
Db 558 GAGTAGGTTAGGATAGAAA 541

RESULT 14

US-10-972-079-67863/c
; Sequence 67863, Application US/10972079
; Publication No. US2005015317A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67863
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894338179_2
US-10-972-079-67863

Query Match 82.0%; Score 16.4; DB 9; Length 600;
Best Local Similarity 94.4%; Pred. No. 6.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAA 18
|||||
Db 386 GAGTAGGTTAGGATAGAAA 369

RESULT 15

US-09-925-065A-70545/c
; Sequence 70545, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 70545
; LENGTH: 770
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-925-065A-70545

Query Match 82.0%; Score 16.4; DB 4; Length 770;
Best Local Similarity 94.4%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAA 18
|||||
Db 189 GAGTAGGAAGGAAGAAA 172

Search completed: January 2, 2006, 15:30:57
Job time : 452 secs

THIS PAGE IS BLANK

THIS PAGE IS BLANK

~~RESTRICTED~~
~~CONFIDENTIAL~~

179

175 060

STIC-Biotech/ChemLib

From: Yu, Misook
Sent: Tuesday, December 27, 2005 7:09 AM
To: STIC-Biotech/ChemLib
Subject: 09/675,650

Pls search SEQ ID NO: 4 (20 nts primer)

Examiner Misook Yu, Ph.D.
571-272-0839 (Phone)
Art Unit 1642
REM-3A18 (Room)
REM-3C18 (Mail Box)

RECEIVED
DEC 27 2005
STIC/CHEM. DIVISION

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE IS BLANK